

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
27 June 2002 (27.06.2002)

PCT

(10) International Publication Number
WO 02/50289 A1(51) International Patent Classification⁷: C12N 15/79

(21) International Application Number: PCT/US01/50240

(22) International Filing Date:
19 December 2001 (19.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/742,900 19 December 2000 (19.12.2000) US
60/302,885 5 July 2001 (05.07.2001) US
10/006,038 4 December 2001 (04.12.2001) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	09/742,900 (CIP)
Filed on	19 December 2000 (19.12.2000)
US	60/302,885 (CIP)
Filed on	5 July 2001 (05.07.2001)
US	10/006,038 (CIP)
Filed on	4 December 2001 (04.12.2001)

(71) Applicants (for all designated States except US): SEM-BIOSYS GENETICS, INC. [CA/CA]; 110, 2985 - 23rd Avenue, N.E., Calgary, Alberta T1Y 7L3 (CA). SYN-GENTA PARTICIPATIONS AG [CH/CH]; Schwarzwaldallee 215, CH-4058 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants (for US only): VAN ROOIJEN, Gijs [NL/CA]; 3223 Bearspaw Drive N.W., Calgary, Alberta T2L 1T1 (CA). DECKERS, Harm [NL/CA]; 519, 18th Avenue N.W., Calgary, Alberta T2M 0T7 (CA). HEIFETZ, Peter, Bernard [US/US]; 10805 Birch Bluff Avenue, San Diego, CA 92131 (US). BRIGGS, Steven, P. [US/US]; 2454 Mango Way, DelMar, CA 92014 (US). DALMIA, Bipin, Kumar [IN/US]; 7353 Mannix Court, San Diego, CA 92122 (US). DEL VAL, Gregg [CH/US]; 5727 Erlanger Street, San Diego, CA 92122 (US). ZAPLACHINSKI, Steve [CA/US]; 5840 Dalmead Cres., Calgary, Alberta T3A 1E6 (CA). MOLONEY, Maurice [IE/US]; 34 Edgebrook Cove N.W., Calgary, Alberta T3A 5N5 (CA).

(74) Agents: RAMOS, Robert, T. et al.; Campbell & Flores, LLP, 4350 La Jolla Village Drive, 6th Floor, San Diego, CA 92122-1246 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)
- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY,

[Continued on next page]

(54) Title: METHODS FOR THE PRODUCTION OF MULTIMERIC PROTEINS, AND RELATED COMPOSITIONS

(57) Abstract: Improved methods for the production of multimeric-protein-complexes, such as redox proteins and immunoglobulins, in association with oil bodies are described. The redox protein is enzymatically active when prepared in association with the oil bodies. Also provided are related nucleic acids, proteins, cells, plants, and compositions.

WO 02/50289 A1



BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

-1-

METHODS FOR THE PRODUCTION OF MULTIMERIC PROTEINS, AND RELATED COMPOSITIONS

RELATED APPLICATIONS

Benefit of priority under 35 U.S.C. §119(e) is claimed to U.S. provisional
5 application Serial No. 60/302,885, filed July 5, 2001, to van Rooijen, *et al.*,
entitled "METHODS FOR THE PRODUCTION OF REDOX PROTEINS". This
application is also a continuation-in-part of U.S. utility application Serial No.
10/006,038, filed December 4, 2001 to van Rooijen, *et al.*, entitled "METHODS
FOR THE PRODUCTION OF REDOX PROTEINS"; which is a continuation-in-part
10 of U.S. utility application Serial No. 09/742,900, filed December 19, 2000 to
Heifetz, *et al.*, entitled "METHOD OF PRODUCTION AND DELIVERY OF
THIOREDOXIN". This application is also a continuation-in-part of U.S. utility
application Serial No. 09/742,900. The subject matter of each of the provisional
and utility applications is incorporated herein by reference in its entirety.

15 Field Of The Invention

The present invention relates to multimeric-protein-complexes, redox
proteins, and recombinant polypeptides; and improved methods for their
production.

BACKGROUND

20 Multimeric proteins (i.e. proteins comprising multiple polypeptide chains)
are a biologically and commercially important class of proteins. Antibodies for
example are multimeric proteins which are used to treat a wide range of disease
conditions. However in view of their complexity, multimeric proteins frequently
represent significant manufacturing challenges.

25 Redox proteins are also a commercially important class of proteins with
applications in a variety of different industries including the pharmaceutical,
personal care and food industry. For example, the redox protein thioredoxin may
be used in the manufacture of personal care products (Japanese Patent
Applications JP9012471A2, JP103743A2, JP1129785A2), pharmaceutical
30 compositions/products (Aota et al. (1996) J. Cardio. Pharmacol. (1996) 27:
727-732) as well as to reduce protein allergens present in food products such as

-2-

milk (del Val et al. (1999) J. Allerg. Clin. Immunol. 103: 690-697) and wheat (Buchanan et al. (1997) Proc. Natl. Acad. Sci. USA 94: 5372-5377).

However, there is a need in the art to further improve the methods for the recombinant expression of multimeric proteins, including redox proteins. The present invention satisfies this need and provides related advantages as well.

SUMMARY OF THE INVENTION

The present invention relates to novel and improved methods of producing a first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulin-polypeptide-chains, immunoglobulins, redox-fusion-polypeptides, and/or thioredoxin-related proteins; in association with oil bodies. Accordingly, provided herein are methods of producing a recombinant multimeric-protein-complex, said method comprising: (a) producing in a cell comprising oil bodies, a first recombinant polypeptide and a second recombinant polypeptide wherein said first recombinant polypeptide is capable of associating with said second recombinant polypeptide to form said multimeric-protein-complex; and (b) associating said multimeric-protein-complex with an oil body through an oil-body-targeting-protein capable of associating with said oil bodies and said first recombinant polypeptide.

The method further contemplates isolating the oil bodies associated with said recombinant multimeric-protein-complex. The second recombinant polypeptide can be associated with a second oil-body-targeting-protein capable of associating with an oil body and said second recombinant polypeptide. Each of said oil-body-targeting-proteins can be an oil-body-protein or an immunoglobulin. The oil-body-targeting-protein can be an oleosin or caleosin. When the oil-body-targeting-protein can be an oleosin or caleosin, the first recombinant polypeptide can be fused to said oleosin or caleosin. Likewise, the second recombinant polypeptide can be fused to a second oleosin or second caleosin capable of associating with an oil body. The first and second recombinant polypeptides can be produced as a multimeric-fusion-protein comprising said first and second polypeptide, and can form a multimeric-protein-complex. The multimeric-protein-complex can be a heteromultimeric-protein-

-3-

- complex, and the heteromultimeric-protein-complex can be an enzymatically active redox complex or an immunoglobulin. In one embodiment, the first recombinant polypeptide is capable of associating with said second recombinant polypeptide in the cell. In another embodiment, the first recombinant
- 5 polypeptide can be a thioredoxin and the second recombinant polypeptide can be a thioredoxin-reductase. In particular embodiments, the thioredoxin can be selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194; and the thioredoxin-reductase can be selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID
- 10 NOs:195-313. In another embodiment, the first recombinant polypeptide can be an immunoglobulin-polypeptide-chain. For example, the first recombinant polypeptide can be an immunoglobulin light chain, or an immunologically active portion thereof, and the second recombinant polypeptide can be an immunoglobulin heavy chain, or an immunologically active portion thereof. In
- 15 this embodiment, the oil-body-targeting-protein can comprise protein A, protein L or protein G. The cell can be a plant cell, such as a safflower cell, and the like.

Also provided herein is a method of expressing a recombinant multimeric-protein-complex comprising a first and second recombinant polypeptide in a cell, said method comprising:

- 20 (a) introducing into a cell a first chimeric nucleic acid sequence comprising:
- (i) a first nucleic acid sequence capable of regulating transcription in said cell operatively linked to;
 - (ii) a second nucleic acid sequence encoding a first recombinant polypeptide;
- 25 (b) introducing into said cell a second chimeric nucleic acid sequence comprising:
- (i) a third nucleic acid sequence capable of regulating transcription in said cell operatively linked to;
 - (ii) a fourth nucleic acid sequence encoding a second recombinant polypeptide;
- 30 (c) growing said cell under conditions to permit expression of said first and second recombinant polypeptide in a progeny cell comprising oil bodies wherein

-4-

said first recombinant polypeptide and said second recombinant polypeptide are capable of forming a multimeric-protein-complex; and

(d) associating said first recombinant polypeptide with an oil body through an oil-body-targeting-protein capable of associating with said oil bodies and said first

5 recombinant polypeptide. This method further contemplates isolating from the progeny cell, oil bodies comprising the multimeric-protein-complex. The second recombinant polypeptide can be associated with a second oil-body-targeting-protein capable of associating with an oil body and second recombinant polypeptide. Each of said oil-body-targeting-proteins can be an oil-body-protein

10 or an immunoglobulin. The oil-body-targeting-protein can be an oleosin or caleosin. When the oil-body-targeting-protein is an oleosin or caleosin, the first recombinant polypeptide can be fused to said oleosin or caleosin. Likewise, the second recombinant polypeptide can be fused to a second oleosin or second caleosin capable of associating with an oil body. The first and second

15 recombinant polypeptides can be produced as a multimereic-fusion-protein comprising said first and second polypeptide, and can form a multimeric-protein-complex. The multimeric-protein-complex can be a heteromultimeric-protein-complex, and the heteromultimeric-protein-complex can be an enzymatically active redox complex or an immunoglobulin. In one embodiment, the first

20 recombinant polypeptide and said second recombinant polypeptide are capable of forming a multimeric-protein-complex in said progeny cell. In another embodiment, the first recombinant polypeptide can be a thioredoxin and the second recombinant polypeptide can be a thioredoxin-reductase. In particular embodiments, the thioredoxin can be selected from the group consisting of SEQ

25 ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194; and the thioredoxin-reductase can be selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313. In another embodiment, the first recombinant polypeptide can be an immunoglobulin-polypeptide-chain. For example, the first recombinant polypeptide can be an immunoglobulin light chain,

30 or an immunologically active portion thereof, and the second recombinant polypeptide can be an immunoglobulin heavy chain, or an immunologically active portion thereof. In this embodiment, the oil-body-targeting-protein can comprise

-5-

protein A, protein L or protein G. The cell can be a plant cell, such as a safflower cell, and the like.

Also provided herein are methods of producing in a plant a recombinant multimeric-protein-complex, said method comprising:

- 5 (a) preparing a first plant comprising cells, said cells comprising oil bodies and a first recombinant polypeptide wherein said first recombinant polypeptide is capable of associating with said oil bodies through an oil-body-targeting-protein;
- (b) preparing a second plant comprising cells, said cells comprising oil bodies and a second recombinant polypeptide; and
- 10 (c) sexually crossing said first plant with said second plant to produce a progeny plant comprising cells, said cells comprising oil bodies, wherein said oil bodies are capable of associating with said first recombinant polypeptide, and said first recombinant recombinant polypeptide is capable of associating with said second recombinant polypeptide to form said recombinant multimeric-protein-complex.
- 15 The second recombinant polypeptide can be associated with oil bodies through a second oil-body-targeting-protein in the second plant. The oil bodies can be isolated from the progeny plant comprising said multimeric-protein-complex. The oil-body-targeting-protein can be selected from an oil-body-protein or an immunoglobulin, wherein the oil-body-protein can be an oleosin or caleosin. The
- 20 first recombinant polypeptide can be fused to the oleosin or caleosin; and the second recombinant polypeptide can be fused to a second oleosin or second caleosin capable of associating with an oil body. The first and second recombinant polypeptide can form a multimeric-protein-complex, such as a heteromultimeric-protein-complex, wherein the heteromultimeric-protein-complex
- 25 can be an enzymatically active redox complex or an immunoglobulin. In a particular embodiment, the first recombinant polypeptide is a thioredoxin and the second recombinant polypeptide is a thioredoxin-reductase. The thioredoxin can be selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194; and the thioredoxin-reductase can be selected from the group
- 30 consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313. In another embodiment, the first recombinant polypeptide can be an immunoglobulin-polypeptide-chain. For example, the first recombinant

-6-

polypeptide can be an immunoglobulin light chain, or an immunologically active portion thereof, and the second recombinant polypeptide can be an immunoglobulin heavy chain, or an immunologically active portion thereof. In this embodiment, the oil-body-targeting-protein can comprise protein A, protein L
5 or protein G. The plant can be a safflower plant.

Also provided herein are chimeric nucleic acids encoding a multimeric-fusion-protein as described herein, said nucleic acid comprising:

- (a) a first nucleic acid sequence encoding an oil-body-targeting-protein operatively linked in reading frame to;
- 10 (b) a second nucleic acid sequence encoding a first recombinant polypeptide; linked in reading frame to;
- (c) a third nucleic acid sequence encoding a second recombinant polypeptide, wherein said first and second recombinant polypeptide are capable of forming a multimeric-protein-complex. The oil-body-targeting-protein can be selected from
15 an oil-body-protein or an immunoglobulin. The oil-body-protein can be an oleosin or caleosin. The multimeric-protein-complex can be a heteromultimeric-protein-complex, and the first and second recombinant polypeptide can form an enzymatically active heteromultimeric redox complex or an immunoglobulin. In a particular embodiment, the first recombinant polypeptide is a thioredoxin and the
20 second recombinant polypeptide is a thioredoxin-reductase. The thioredoxin can be selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194; and the thioredoxin-reductase can be selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313. In another embodiment, the first recombinant polypeptide can be
25 an immunoglobulin-polypeptide-chain. For example, the first recombinant polypeptide can be an immunoglobulin light chain, or an immunologically active portion thereof, and the second recombinant polypeptide can be an immunoglobulin heavy chain, or an immunologically active portion thereof. In this embodiment, the oil-body-targeting-protein can comprise protein A, protein L
30 or protein G. In yet another embodiment, positioned between the nucleic acid sequence encoding an oil-body-targeting-protein and the nucleic acid sequence encoding a first recombinant polypeptide can be a linker nucleic acid sequence

-7-

encoding an oil-body-surface-avoiding linker amino acid sequence. The oil-body-surface-avoiding linker amino acid sequence can be substantially negatively charged, or have a molecular weight of at least 35 kd. Optionally, the gene fusion further comprises a linker nucleic acid sequence encoding an amino acid
5 sequence that is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is positioned between the oil-body-surface-avoiding linker amino acid sequence that is also a non-proteolytic linker and said sequence encoding the first recombinant polypeptide.

Also provided herein are recombinant multimeric-fusion-proteins
10 comprising (i) an oil-body-targeting-protein, or fragment thereof, (ii) a first recombinant polypeptide and a (iii) second recombinant polypeptide, wherein said first and second recombinant polypeptides are capable of forming a multimeric-protein-complex. The oil-body-targeting-protein can be selected from an oil-body-protein or an immunoglobulin, and the oil-body-protein can be an
15 oleosin or a caleosin. The multimeric-fusion-protein can be a heteromultimeric-fusion-protein, wherein said first and second recombinant polypeptide form an enzymatically active heteromultimeric redox complex or an immunoglobulin. In a particular embodiment, the first recombinant polypeptide is a thioredoxin and the second recombinant polypeptide is a thioredoxin-reductase. The thioredoxin can
20 be selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194; and the thioredoxin-reductase can be selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313. In another embodiment, the first recombinant polypeptide can be an immunoglobulin-polypeptide-chain. For example, the first recombinant
25 polypeptide can be an immunoglobulin light chain, or an immunologically active portion thereof, and the second recombinant polypeptide can be an immunoglobulin heavy chain, or an immunologically active portion thereof. In this embodiment, the oil-body-targeting-protein can comprise protein A, protein L or protein G. In yet another embodiment, positioned between the nucleic acid
30 sequence encoding an oil-body-targeting-protein and the nucleic acid sequence encoding a first recombinant polypeptide can be a linker nucleic acid sequence encoding an oil-body-surface-avoiding linker amino acid sequence. The oil-body-

surface-avoiding linker amino acid sequence can be substantially negatively charged, or have a molecular weight of at least 35 kd. Optionally, the gene fusion further comprises a linker nucleic acid sequence encoding an amino acid sequence that is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is positioned between the oil-body-surface-avoiding linker amino acid sequence and said sequence encoding the first recombinant polypeptide.

Also provided herein are isolated oil bodies comprising a multimeric-protein-complex comprising (i) an oil-body-targeting-protein and (ii) a first recombinant polypeptide, said oil bodies further comprising a second recombinant polypeptide, wherein said first and second recombinant polypeptide are capable of forming a multimeric-protein-complex. The oil-body-targeting-protein can be selected from an oil-body-protein or an immunoglobulin, and the oil-body-protein can be an oleosin or a caleosin. The multimeric-fusion-protein can be a heteromultimeric-fusion-protein, wherein said first and second recombinant polypeptide form an enzymatically active heteromultimeric redox complex or an immunoglobulin. In a particular embodiment, the first recombinant polypeptide is a thioredoxin and the second recombinant polypeptide is a thioredoxin-reductase. In another embodiment, the first recombinant polypeptide can be an immunoglobulin-polypeptide-chain. For example, the first recombinant polypeptide can be an immunoglobulin light chain, or an immunologically active portion thereof, and the second recombinant polypeptide can be an immunoglobulin heavy chain, or an immunologically active portion thereof. In this embodiment, the oil-body-targeting-protein can comprise protein A, protein L or protein G.

Also provided herein are isolated oil bodies comprising (a) a first fusion protein comprising a first oil-body-targeting-protein fused to a first recombinant polypeptide; and (b) a second fusion protein comprising a second oil-body-targeting-protein fused to a second recombinant polypeptide, wherein said first and second recombinant polypeptide are capable of forming a multimeric-protein-complex. The oil-body-targeting-protein can be selected from an oil-body-protein or an immunoglobulin, and the oil-body-protein can be an

oleosin or a caleosin. The multimeric-fusion-protein can be a heteromultimeric-fusion-protein, wherein said first and second recombinant polypeptide form an enzymatically active heteromultimeric redox complex or an immunoglobulin. In a particular embodiment, the first recombinant polypeptide is a thioredoxin and the
5 second recombinant polypeptide is a thioredoxin-reductase. The thioredoxin can be selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194; and the thioredoxin-reductase can be selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313. In another embodiment, the first recombinant polypeptide can be
10 an immunoglobulin-polypeptide-chain. For example, the first recombinant polypeptide can be an immunoglobulin light chain, or an immunologically active portion thereof, and the second recombinant polypeptide can be an immunoglobulin heavy chain, or an immunologically active portion thereof. In this embodiment, the oil-body-targeting-protein can comprise protein A, protein L
15 or protein G.

Also provided are cells and transgenic plants comprising oil bodies, multimeric-protein-complexes, and multimeric-fusion-proteins, set forth herein. In one embodiment, the first recombinant polypeptide can be an immunoglobulin-polypeptide-chain. For example, the first recombinant polypeptide can be an
20 immunoglobulin light chain, or an immunologically active portion thereof, and the second recombinant polypeptide can be an immunoglobulin heavy chain, or an immunologically active portion thereof. In this embodiment, the oil-body-targeting-protein can comprise protein A, protein L or protein G. In
embodiments, wherein said first recombinant polypeptide is a thioredoxin and
25 said second recombinant polypeptide is a thioredoxin-reductase, the methods described herein can be used to formulate the oil bodies for use in the preparation of a food product, personal care product or pharmaceutical composition. These formulations can further comprise the addition of NADP or NADPH. The food product can be a milk or wheat based food product. The
30 personal care product can reduce the oxidative stress to the surface area of the human body or can be used to lighten the skin. The pharmaceutical composition can be used to treat chronic obstructive pulmonary disease (COPD), cataracts,

-10-

diabetes, envenomation, bronchiopulmonary disease, malignancies, psoriasis, reperfusion injury, wound healing, sepsis, GI bleeding, intestinal bowel disease (IBD), ulcers, GERD (gastro esophageal reflux disease).

Also provided herein are compositions comprising isolated oil bodies,
5 thioredoxin and thioredoxin-reductase, wherein said thioredoxin can be selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194, and said thioredoxin-reductase can be selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313. The composition can further comprise NADP or NADPH. In another
10 embodiment, the composition comprises a first recombinant polypeptide that can be an immunoglobulin-polypeptide-chain and a second recombinant polypeptide. For example, the first recombinant polypeptide can be an immunoglobulin light chain, or an immunologically active portion thereof, and the second recombinant polypeptide can be an immunoglobulin heavy chain, or an immunologically active
15 portion thereof. In this embodiment, the oil-body-targeting-protein can comprise protein A, protein L or protein G.

Also provided are multimeric-fusion-proteins, wherein the fusion-protein contains two or more polypeptide chains selected from the group of proteins set forth in Figure 5. Methods are also provided of reducing allergenicity of a food
20 comprising the steps of providing the isolated oil bodies set forth herein; and adding the isolated oil bodies to the food, whereby allergenicity of the food is reduced. The food can be selected from the group consisting of wheat flour, wheat dough, milk, cheese, yogurt and ice cream. The various methods of treating food can further comprise providing NADH as a co-factor in the
25 substantial absence of NADPH.

Also provided herein are methods of treating or protecting a target against oxidative stress, comprising the steps of providing the recombinant redox fusion polypeptide comprising thioredoxin and thioredoxin-reductase; and contacting the recombinant fusion polypeptide with a target, wherein the target
30 is susceptible to oxidative stress, thereby treating or protecting against the stress. The target can be selected from the group consisting of a molecule, a molecular complex, a cell, a tissue, and an organ.

-11-

Also provided herein are methods for preparing an enzymatically active redox protein associated with oil bodies comprising:

- a) producing in a cell a redox fusion polypeptide comprising a first redox protein linked to a second redox protein;
- 5 b) associating said redox fusion polypeptide with oil bodies through an oil-body-targeting-protein capable of associating with said redox fusion polypeptide and said oil bodies; and
- 10 c) isolating said oil bodies associated with said redox fusion polypeptide. The first redox protein can be a thioredoxin and the second redox protein can be a thioredoxin-reductase.

Also, provided herein are methods of producing an immunoglobulin, said method comprising: (a) producing in a cell comprising oil bodies, a first immunoglobulin-polypeptide-chain and a second immunoglobulin-polypeptide-chain wherein said first immunoglobulin-polypeptide-chain is capable of
15 associating with said second immunoglobulin-polypeptide-chain to form said immunoglobulin; and (b) associating said immunoglobulin with an oil body through an oil-body-targeting-protein capable of associating with said oil bodies and said first immunoglobulin-polypeptide-chain. For example, the first immunoglobulin-polypeptide-chain can be an immunoglobulin light chain, or an
20 immunologically active portion thereof, and the second immunoglobulin-polypeptide-chain can be an immunoglobulin heavy chain, or an immunologically active portion thereof. In this embodiment, the oil-body-targeting-protein can comprise protein A, protein L or protein G.

Also provided herein are methods for preparing a redox protein or an
25 immunoglobulin associated with oil bodies comprising:

- a) introducing into a cell a chimeric nucleic acid sequence comprising:
 - 1) a first nucleic acid sequence capable of regulating transcription in said cell operatively linked to;
 - 2) a second nucleic acid sequence encoding a recombinant
30 fusion polypeptide comprising (i) a nucleic acid sequence encoding a sufficient portion of an oil-body-protein to provide targeting of said recombinant fusion polypeptide to an oil body linked to (ii) a

-12-

- nucleic acid sequence encoding a redox fusion polypeptide comprising a first redox protein linked to a second redox protein, or a nucleic acid sequence encoding an immunoglobulin comprising a first immunoglobulin-polypeptide-chain linked to a second immunoglobulin-polypeptide-chain, operatively linked to;
- 5 3) a third nucleic acid sequence capable of terminating transcription in said cell;
- b) growing said cell under conditions to permit expression of said redox fusion polypeptide or immunoglobulin in a progeny cell comprising oil
- 10 bodies; and
- c) isolating from said progeny cell said oil bodies comprising said redox fusion polypeptide or immunoglobulin. In certain embodiments, positioned between said nucleic acid sequence encoding a sufficient portion of an oil-body-protein and said nucleic acid sequence encoding a redox fusion polypeptide or
- 15 immunoglobulin can be a linker nucleic acid sequence encoding an oil-body-surface-avoiding linker amino acid sequence. The oil-body-surface-avoiding linker amino acid sequence can be substantially negatively charged or have a molecular weight of at least 35 kd. Optionally, the gene fusion further comprises a linker nucleic acid sequence encoding an amino acid sequence that
- 20 is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is positioned between the oil-body-surface-avoiding linker amino acid sequence and said nucleic acid sequence encoding a redox fusion polypeptide. In this optional embodiment, also contemplated is the introduction of an enzyme or chemical that cleaves said redox fusion polypeptide from said oil body, thereby
- 25 obtaining isolated redox fusion polypeptide. The first redox protein can be a thioredoxin and said second redox protein can be a thioredoxin-reductase. In one embodiment, the thioredoxin and thioredoxin-reductase can be obtained from *Arabidopsis*. In another embodiment, the first redox protein is at least 5 times more active when produced as a redox fusion polypeptide as compared to
- 30 the production of the first redox protein without the second redox protein.

Also provided herein, for use with the various methods set forth herein is the formulation of an emulsion of the oil bodies associated with the redox fusion

-13-

polypeptide for use in the preparation of a product capable of treating oxidative stress in a target, a product capable of chemically reducing a target, pharmaceutical composition, a personal care product or a food product. Accordingly, an emulsion formulation composition is provided.

5 Also provided herein is a chimeric nucleic acid comprising:

- 1) a first nucleic acid sequence capable of regulating transcription in a host cell operatively linked to;
- 2) a second nucleic acid sequence encoding a recombinant fusion polypeptide comprising (i) a nucleic acid sequence encoding a sufficient portion
10 of an oil-body-protein to provide targeting of said recombinant fusion polypeptide to an oil body linked to (ii) a nucleic acid sequence encoding a redox fusion polypeptide comprising a first redox protein linked to a second redox protein operatively linked to;
- 3) a third nucleic acid sequence capable of terminating transcription
15 in said cell. The oil-body-protein can be an oleosin or a caleosin, the first redox protein can be a thioredoxin and said second redox protein can be a thioredoxin-reductase. In certain embodiments, positioned between said nucleic acid sequence encoding a sufficient portion of an oil-body-protein and said nucleic acid sequence encoding a redox fusion polypeptide is a linker nucleic acid
20 sequence encoding an oil-body-surface-avoiding linker amino acid sequence. The oil-body-surface-avoiding linker amino acid sequence can be substantially negatively charged, or have a molecular weight of at least 35 kd. In one embodiment, the gene fusion optionally further comprises a linker nucleic acid sequence encoding an amino acid sequence that is specifically cleavable by an
25 enzyme or a chemical, wherein the linker sequence is positioned between the oil-body-surface-avoiding linker amino acid sequence and said nucleic acid sequence encoding a redox fusion polypeptide.

Also provided herein are transgenic plants, e.g., safflower plants, comprising any of the chimeric nucleic acid sequences and constructs described
30 herein. The chimeric nucleic acids can be contained within a plastid. Accordingly, isolated plastids are provided having chimeric nucleic acids therein.

-14-

Also provided are plant seeds comprising the chimeric nucleic acids provided herein.

Also provided are oil body preparations obtained using any of the methods provided herein, and food products, pharmaceutical compositions, and personal care products containing the oil body preparations. The products and/or compositions provided herein are capable of treating oxidative stress in a target, capable of chemically reducing a target. Also provided is a detergent composition comprising an oil body preparation capable of chemically reducing a target, and related methods of cleansing an item, comprising administering such product to the item under conditions that promote cleansing.

Also provided herein are nucleic acid constructs comprising a gene fusion, wherein the gene fusion comprises a first region encoding an oil-body-protein or an active fragment thereof, operably linked to a second region encoding at least one thioredoxin-related protein or an active fragment thereof. In one embodiment, the at least one thioredoxin-related protein can be thioredoxin. The thioredoxin can be selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194. The thioredoxin can be obtained from *Arabidopsis* or wheat.

In another embodiment, the at least one thioredoxin-related protein can be thioredoxin-reductase. The thioredoxin-reductase can be selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313 and/or derived from *Arabidopsis* or wheat. The thioredoxin-reductase can be an NADPH-dependent thioredoxin-reductase. The second region can encode a thioredoxin and thioredoxin-reductase. In one embodiment, the thioredoxin and thioredoxin-reductase is obtained from *Mycobacterium leprae*. In another embodiment, the at least one thioredoxin-related protein can be an engineered fusion protein. The first region can precede, in a 5' to 3' direction, the second region. Alternatively, the first region follows, in a 5' to 3' direction, the second region. The gene fusion can optionally further comprise a third region encoding a second thioredoxin-related protein or an active fragment thereof, operably linked to the first region, or to the second region, or to both. A seed-specific promoter, such as a phaseolin promoter, can be operably linked to

-15-

the gene fusion. In one embodiment, at least one thioredoxin-related protein is derived from a plant species selected from the group consisting of *Arabidopsis* and wheat. In another embodiment, at least one thioredoxin-related protein can be derived from *E. coli*.

5 In one embodiment, the gene fusion further comprises a nucleic acid sequence encoding an oil-body-surface-avoiding linker amino acid sequence, wherein the linker amino acid sequence is positioned between the first region and the second region. The oil-body-surface-avoiding linker amino acid sequence can be substantially negatively charged, or have a molecular weight of at least
10 35 kd. In addition, the gene fusion can further comprise a linker nucleic acid sequence encoding an amino acid sequence that is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is positioned between the oil-body-surface-avoiding linker amino acid sequence and the second region.

 Also provided herein are transgenic plants containing a nucleic acid
15 construct comprising a gene fusion, wherein the gene fusion comprises a region encoding an oil-body-protein or an active fragment thereof, operably linked to a region encoding a first thioredoxin-related protein or an active fragment thereof. The thioredoxin-related protein can be thioredoxin. The nucleic acid construct can be contained within a plastid. In one embodiment, when the first
20 thioredoxin-related protein is thioredoxin and the construct can further comprise a region encoding a thioredoxin-reductase. The gene fusion can optionally further comprise a third region encoding a second thioredoxin-related protein or an active fragment thereof, operably linked to the first region, or to the second region, or to both. The gene fusion can optionally further comprise a nucleic
25 acid sequence encoding an oil-body-surface-avoiding linker amino acid sequence, wherein the nucleic acid encoding the linker amino acid sequence is positioned between the region encoding an oil-body-protein and the region encoding a first thioredoxin-related protein. The oil-body-surface-avoiding linker amino acid sequence can be substantially negatively charged, or have a molecular weight of
30 at least 35 kd. The gene fusion can optionally further comprise a linker nucleic acid sequence encoding an amino acid sequence that is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is

-16-

positioned between the oil-body-surface-avoiding linker amino acid sequence and the region encoding a first thioredoxin-related protein.

Also provided is a transgenic plant comprising a nucleic acid construct, a seed-specific promoter operably linked to a gene fusion, wherein the gene fusion
5 comprises a region encoding an oil-body-protein or an active fragment thereof, operably linked to a region encoding a first thioredoxin-related protein or an active fragment thereof, wherein a fusion protein comprising activities of oleosin and the thioredoxin-related protein is produced in a seed of the plant. In another embodiment, a thioredoxin-related protein having concentration of at least about
10 0.5% of total cellular seed protein is provided. Also provided herein is an extract comprising an activity of a thioredoxin-related protein. Also provided are oil bodies and/or oil obtained from various seeds.

Also provided herein are methods of making a fusion protein comprising a thioredoxin-related activity, the method comprising the steps of:

- 15 a) providing a transgenic plant comprising a nucleic acid construct comprising a seed-specific promoter operably linked to a gene fusion, wherein the gene fusion comprises a region encoding an oil-body-protein or an active fragment thereof, operably linked to a region encoding a first thioredoxin-related protein or an active fragment thereof, the gene fusion encoding a fusion protein
20 comprising a thioredoxin-related activity;
- b) obtaining seeds from the plant; and
- c) recovering the fusion protein by isolating oil bodies from the seeds. In one embodiment, the oil bodies are fractionated to achieve partial purification of the fusion protein. The oil bodies can be in association with a fusion protein.
- 25 The oil-body-protein can be cleaved from the thioredoxin-related protein after fractionation of the oil bodies. The cleaving step can make use of a protease or chemical proteolysis.

Also provided herein are methods of reducing allergenicity of a food comprising the steps of:

- 30 a) providing a preparation comprising oil bodies associated with a fusion protein, the fusion protein comprising an oil-body-protein or an active fragment thereof and a thioredoxin-related protein or an active fragment thereof; and

-17-

b) adding the preparation to the food, whereby allergenicity of the food is reduced due to activity of the thioredoxin-related protein or fragment. The food can be wheat flour, wheat dough, milk, cheese, yogurt and ice cream. In one embodiment, NADH is used as a co-factor in the substantial absence of NADPH.

5 Also provided herein are pharmaceutical compositions comprising a fusion protein, the fusion protein comprising an oil-body-protein or an active fragment thereof and a thioredoxin-related protein or an active fragment thereof, in a pharmaceutically acceptable carrier. The oil bodies can be associated with the fusion protein. Also provided is a cosmetic formulation comprising oil bodies
10 associated with a fusion protein, the fusion protein comprising an oil-body-protein or an active fragment thereof and a thioredoxin-related protein or an active fragment thereof, in a pharmaceutically acceptable carrier. Also provided are methods of treating or protecting a target against oxidative stress, comprising the steps of:

15 a) providing a preparation comprising a fusion protein, the fusion protein comprising an oil-body-protein or an active fragment thereof and a thioredoxin-related protein or an active fragment thereof; and

b) contacting the preparation with a target, wherein the target is susceptible to oxidative stress, thereby treating or protecting against the stress.

20 The target can be selected from the group consisting of a molecule, a molecular complex, a cell, a tissue, and an organ.

Also provided is a nucleic acid construct comprising a gene fusion, wherein the gene fusion comprises a first region encoding an oil-body-protein or an active fragment thereof, operably linked to a second region encoding at least
25 one polypeptide or an active fragment thereof, and an oil-body-surface-avoiding linker in frame between the first and second region polypeptides. Also provided are methods of expressing this construct into the encoded amino acid sequence; and oil bodies, formulations, emulsions, cells, and plants comprising the construct and encoded amino acid sequence. These particular constructs, oil
30 bodies, formulations, emulsions, cells, and plants can be produced according to the methods described herein. The second region can encode any polypeptide, for example, a therapeutically, nutritionally, industrially or cosmetically useful

-18-

peptide as set forth herein. For example, the second region can encode a redox protein, an immunoglobulin, a thioredoxin-related protein or any one or more recombinant polypeptides of a multimeric-protein-complex.

Other features and advantages of the present invention will become
5 readily apparent from the following detailed description. It should be understood however that the detailed description and the specific examples while indicating particular embodiments of the invention are given by way of illustration only.

BRIEF DESCRIPTION OF THE DRAWINGS

10 Figure 1 shows a ClustalW Formatted Alignment comparison of the published NADPH thioredoxin-reductase nucleic acid sequence (SEQ ID NO:9) (ATTHIREDB-Jacquot et al. J. Mol. Biol. (1994) 235 (4):1357-63.) with the sequence isolated herein in Example 1 (TR; SEQ ID NO:8).

Figure 2 shows a ClustalW Formatted Alignment comparison of the
15 deduced amino acid sequence of the published NADPH thioredoxin-reductase sequence (SEQ ID NO:12)(ATTHIREDB Jacquot et al. J. Mol. Biol. (1994) 235 (4):1357-63.) with the sequence isolated herein in Example 1 (TR; SEQ ID NO:13).

Figure 3 shows a clustal alignment comparing the amino acid sequence of
20 the *Arabidopsis thaliana* thioredoxin-reductase-linker-thioredoxin synthetic fusion (Arab TR-link-Trxh; SEQ ID NO:37) to the *Mycobacterium leprae* thioredoxin-reductase-thioredoxin natural fusion (M.lep TR/Trxh; SEQ ID NO:36) natural fusion. Overall, the proteins are approximately 50% identical at the amino acid level.

25 Figure 4 is a bar graph showing the thioredoxin/thioredoxin-reductase activity measurements for the various transgenic *Arabidopsis* seed fractions. Relative specific activity is expressed as a percentage of the *E. coli* thioredoxin and thioredoxin-reductase activities. The numbered bars in the graph correspond to the following:

- 30
1. W.T. + oleosin-thioredoxin
 2. W.T. + thioredoxin-oleosin
 3. W.T. + thioredoxin

-19-

4. W.T. + oleosin-thioredoxin-reductase
5. W.T. + thioredoxin-reductase-oleosin
6. W.T. + thioredoxin-reductase
7. thioredoxin + oleosin-thioredoxin-reductase
- 5 8. thioredoxin + thioredoxin-reductase-oleosin
9. thioredoxin + thioredoxin-reductase
10. thioredoxin-reductase + oleosin-thioredoxin
11. thioredoxin-reductase + thioredoxin-oleosin
12. oleosin-*M.lep* TR/Trxh
- 10 13. *E. coli* thioredoxin-reductase + thioredoxin

Figure 5 provides a listing of exemplary proteins for use in the heteromultimeric-fusion-proteins and heteromultimeric-protein-complexes provided herein.

DETAILED DESCRIPTION

- 15 As hereinbefore mentioned, the present invention relates to novel and improved methods for the production of multimeric proteins, including a first and second recombinant polypeptide, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulin-polypeptide-chains,
- 20 immunoglobulins, redox-fusion-polypeptides, and a first and second thioredoxin-related protein; and related products. These methods permit the production of active multimeric-protein-complexes in association with oil bodies. The oil bodies in association with the multimeric-protein-complex may be used to prepare various useful emulsions.
- 25 Accordingly, provided herein are methods of producing a recombinant multimeric-protein-complex associated with an oil body, said method comprising:
- (a) producing in a cell comprising oil bodies, a first recombinant polypeptide and a second recombinant polypeptide wherein said first recombinant polypeptide is capable of associating with said second recombinant
- 30 polypeptide in the cell to form said multimeric-protein-complex; and

(b) associating said multimeric-protein-complex with an oil body through an oil-body-targeting-protein capable of associating with said oil body and said first recombinant polypeptide.

Definitions and terms

5 Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which this invention belongs. Where permitted, all patents, applications, published applications and other publications and sequences from GenBank, SwissPro and other data bases referred to throughout in the disclosure herein are
10 incorporated by reference in their entirety.

 As used herein, the phrase "multimeric-protein-complex", refers to two or more polypeptide chains that permanently or repeatedly interact or permanently or repeatedly coordinate to form a biologically active assembly comprising said two or more polypeptide chains. It should be noted that the polypeptides may
15 be independently biologically active without interaction or coordination to form the complex. The multimeric-protein-complex may provide a biological structure, or it may be capable of facilitating a chemical or biological reaction. For example, one of the protein regions within the multimeric-protein-complex can repeatedly activate or repeatedly inactivate the biological or metabolic activity of
20 one or more of the other proteins contained within the multimeric-protein-complex. In one embodiment, the first and second recombinant polypeptide contained in a multimeric-protein-complex may either associate or interact as independent non-contiguous polypeptide chains or the multimeric-protein-complex may be prepared as a fusion polypeptide (multimeric-fusion-protein)
25 between the first and second recombinant polypeptide.

 One example of a repeated (e.g., reoccurring) interaction or association between the two or more polypeptides of a multimeric-protein-complex provided herein is the interaction between two or more non-identical redox proteins to form a heteromultimeric-protein-complex. Exemplary redox proteins for use in
30 this regard are thioredoxin and the thioredoxin-reductase. A further example is the interaction between two or more immunoglobulin-polypeptide-chains to form an immunoglobulin. As used herein, the phrase "heteromultimeric-protein-

-21-

complex", refers to two or more non-identical polypeptide chains that permanently or repeatedly interact or permanently or repeatedly coordinate to form a biologically active assembly comprising said two or more polypeptide chains. Other examples of multimeric-protein-complexes provided herein include

5 a first and second recombinant polypeptide, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, first and second immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and a first and second thioredoxin-related protein.

The recombinant polypeptide or multimeric-protein-complex is associated

10 with an oil body. As used herein, the phrase "oil body" or "oil bodies" refers to any oil or fat storage organelle in any cell type. Accordingly, the oil bodies may be obtained from any cell comprising oil bodies, including plant cells (described in for example: Huang (1992) Ann. Rev. Plant Mol. Biol. 43: 177-200), animal cells (described in for example: Murphy (1990) Prog Lipid Res 29(4): 299-324),

15 including adipocytes, hepatocytes, steroidogenic cells, mammary epithelial cells, macrophages, algae cells (described in for example: Rossler (1988) J. Physiol. London, 24: 394-400) fungal cells, including yeast cells (described in for example Leber et al. (1994) Yeast 10: 1421-1428) and bacterial cells (described in for example: Pieper-Furst et al. (1994) J. Bacteriol. 176: 4328-4337).

20 Preferably the oil bodies used herein are oil bodies obtainable from plant cells and more preferably the oil bodies obtainable from plant seed cells.

As used herein, the phrase "is capable of associating with", "associate" or grammatical variations thereof, refers to any interaction between two or more polypeptides, including any covalent interactions (e.g. multimeric-fusion-proteins)

25 as well as non-covalent interactions. Exemplary non-covalent interactions can be between the oil-body-targeting-protein and a redox protein or immunoglobulin-polypeptide-chain, as well as between two or more different proteins contained within two or more separate oil-body-protein fusion proteins (e.g., the redox proteins in oleosin-thioredoxin and oleosin-thioredoxin-reductase).

30 As used herein, the term "recombinant" (also referred to as heterologous) in the context of recombinant proteins and amino acids, means "of different natural origin" or represents a non-natural state. For example, if a host cell is

-22-

transformed with a nucleotide sequence derived from another organism, particularly from another species, that nucleotide sequence and amino acid sequence encoded thereby, is recombinant (heterologous) with respect to that host cell and also with respect to descendants of the host cell which carry that gene. Similarly, recombinant (or heterologous) refers to a nucleotide sequence derived from and inserted into the same natural, original cell type, but which is present in a non-natural state, e.g., a different copy number, or under the control of different regulatory elements. A transforming nucleotide sequence may include a recombinant coding sequence, or recombinant regulatory elements.

Alternatively, the transforming nucleotide sequence may be completely heterologous or may include any possible combination of heterologous and endogenous nucleic acid sequences.

In various embodiments of the present invention, the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and/or thioredoxin-related proteins, are produced in a cell comprising oil bodies. As used herein the phrase "in a cell", "in the cell", or grammatical variations thereof, mean that the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and/or thioredoxin-related proteins, may be produced in any cellular compartment of that cell, so long as that cell comprises oil bodies therein. In embodiments of the invention in which plant cells are used, the phrase is intended to include the plant apoplast.

In various embodiments provided herein, the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and thioredoxin-related proteins, associate with an oil body through an oil-body-targeting-protein. As used herein, the phrase "oil-body-targeting-protein" refers to any protein, protein fragment or peptide capable of associating with an oil

-23-

body. Exemplary oil-body-targeting-proteins for use herein include oil-body-proteins, such as oleosin and caleosin; immunoglobulins, such as bi-specific antibodies; and the like.

In embodiments described herein in which an oil-body-protein is used, the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and thioredoxin-related proteins, are preferably fused to the oil-body-protein. The term "oil-body-protein" refers to any protein naturally present in cells and having the capability of association with oil bodies, including any oleosin or caleosin.

Accordingly, provided herein a method of expressing a recombinant multimeric-protein-complex comprising a first and second recombinant polypeptide in a cell, said method comprising:

- 15 (a) introducing into a cell a first chimeric nucleic acid sequence comprising:
 - (i) a first nucleic acid sequence capable of regulating transcription in said cell operatively linked to;
 - (ii) a second nucleic acid sequence encoding a first recombinant polypeptide, such as a redox protein, an immunoglobulin-polypeptide-chain or an thioredoxin-related protein, fused to an oil-body-protein;
- 20 (b) introducing into said cell a second chimeric nucleic acid sequence comprising:
 - (i) a third nucleic acid sequence capable of regulating transcription in said cell operatively linked to;
 - (ii) a fourth nucleic acid sequence encoding a second recombinant polypeptide, such as a second redox protein, a second immunoglobulin-polypeptide-chain or a second thioredoxin-related protein,;
- 25 (c) growing said cell under conditions to permit expression of said first and second recombinant polypeptide in a progeny cell comprising oil bodies wherein said first recombinant polypeptide and said second recombinant polypeptide are
- 30 capable of forming a multimeric-protein-complex, preferably in said progeny cell; and

-24-

(d) associating said first recombinant polypeptide with an oil body through said oil-body-protein.

The term "nucleic acid" as used herein refers to a sequence of nucleotide or nucleoside monomers consisting of naturally occurring bases, sugars and intersugar (backbone) linkages. The term also includes modified or substituted sequences comprising non-naturally occurring monomers or portions thereof, which function similarly. The nucleic acid sequences may be ribonucleic acids (RNA) or deoxyribonucleic acids (DNA) and may contain naturally occurring bases including adenine, guanine, cytosine, thymidine and uracil. The sequences also may contain modified bases such as xanthine, hypoxanthine, 2-aminoadenine, 6-methyl, 2-propyl and other alkyl adenines, 5-halo-uracil, 5-halo cytosine, 6-aza uracil, 6-aza cytosine and 6-aza thymine, pseudo uracil, 4-thiouracil, 8-halo adenine, 8-amino adenine, 8-thiol-adenine, 8-thio-alkyl adenines, 8-hydroxyl adenine and other 8-substituted adenines, 8-halo guanines, 8 amino guanine, 8 thiol guanine, 8-thioalkyl guanines, 8 hydroxyl guanine and other 8-substituted guanines, other aza and deaza uracils, thymidines, cytosines, adenines, or guanines, 5-trifluoromethyl uracil and 5-trifluoro cytosine.

Multimeric-protein-complexes

In accordance with the methods and compositions provided herein, any two recombinant polypeptides capable of forming a multimeric-protein-complex may be used. The nucleic acid sequences encoding the two recombinant polypeptides may be obtained from any biological source or may be prepared synthetically. In general nucleic acid sequence encoding multimeric proteins are known to the art and readily available. Known nucleic acid sequences encoding multimeric-protein-complexes may be used to design and construct nucleic acid sequence based probes in order to uncover and identify previously undiscovered nucleic acid sequences encoding multimeric-protein-complexes, for example, by screening cDNA or genomic libraries or using 2- or multi-hybrid systems. Thus, additional nucleic acid sequences encoding multimeric-protein-complexes may be discovered and used as described herein.

The first and/or second recombinant polypeptides that are comprised within a multimeric-protein-complex provided herein, can themselves be in the

-25-

form of heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and/or a first and/or second thioredoxin-related protein.

- 5 The nucleic acid sequence encoding the first and second recombinant polypeptide, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and/or a first and/or second thioredoxin-related protein may be obtained from separate sources or may be obtained from
- 10 the same source. In general however, such nucleic acid sequence is obtained from the same or a similar biological source. In certain embodiments wherein the nucleic acid sequence encoding the first and second recombinant polypeptide protein are obtained from the same source, the nucleic acid sequence encoding the first recombinant polypeptide and second recombinant polypeptide may be
- 15 naturally fused. In accordance with a particular embodiment, the nucleic acid sequences encoding the first and second recombinant polypeptide are obtained from a plant source.

Oil-Body-Surface-Avoiding Linkers

- Polypeptide spacers or linkers of variable length and/or negative charge
- 20 can be used herein to separate the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and the first and/or second thioredoxin-related proteins from the in-frame oil-body-targeting-
- 25 protein, to improve activity of and/or the accessibility of the polypeptide or complex. For example, in one embodiment set forth herein, positioned between a nucleic acid sequence encoding a sufficient portion of an oil-body-protein and a nucleic acid sequence encoding either the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes,
- 30 multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and the first

-26-

and/or second thioredoxin-related proteins; is a linker nucleic acid sequence encoding an oil-body-surface-avoiding linker amino acid sequence.

Oil-body-surface-avoiding linkers are positioned between the oil-body targeting sequence and an in-frame recombinant polypeptide of interest, e.g., the multimeric-protein-complexes provided herein, serve to increase the distance and or decrease the interaction between the negatively charged oil body surface and the recombinant polypeptide of interest. A negatively charged linker is repelled by the negatively charged oil body surface, in turn increasing the distance or decreasing the interaction of its attached recombinant polypeptide with the oil body surface. As a consequence of the increased distance from the oil body surface, the recombinant polypeptide will be more accessible, e.g. to its target(s) substrate, protein substrate, protein partner, and less affected by the charged oil body surface. Exemplary linker sequences for use herein can be either a negatively charged linker, or a linker having a molecular weight of at least about 35 kd or more.

As used herein, a "negatively charged linker" sequence, refers to any amino acid segment, or nucleic acid encoding such, that has a pI less than or equal to the pI of an oil body. In certain embodiments, the pI of the negatively charged linker is about 90%, 80%, 70%, 60%, 50%, 40%, 30%, down to about 25% or more, below that of the pI of an oil body in the particular plant or cell system being used. Exemplary negatively charged linkers can be prepared comprising any combination of the negatively charged amino acid residues. For example, in one embodiment, a negatively charged linker comprises either a poly-glutamate or poly-aspartate sequence, or any combination of both amino acid residues. The negatively charged linker is typically at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or more amino acids in length. The negatively charged linkers are preferably non-proteolytic (e.g., non-proteolytic linkers), having no site for efficient proteolysis. When linker size rather than charge is used to minimize interaction of the recombinant polypeptide of interest with the oil body surface, then the linker is non-proteolytic and ranges in molecular weight from about 35 kd up to about 100 kd. The upper size limit is chosen such that the expression of, the activity of,

-27-

the conformation of, and/or the access to target of, the recombinant polypeptide of interest is not significantly affected by the linker.

In certain embodiments, described herein where a non-proteolytic linker amino acid sequence is employed, the gene fusion or protein fusion (multimeric-
5 fusion-protein) can optionally further comprise a linker nucleic or amino acid sequence encoding a sequence that is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is positioned between the non-proteolytic linker sequence and sequence encoding the desired recombinant protein region, e.g., the first and/or second recombinant polypeptides, multimeric-protein-
10 complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins set forth herein. When a cleavable linker sequence is used herein, in a particular embodiment, it is further downstream than the non-proteolytic linker
15 sequence from the oil-body-targeting-protein region of the fusion protein. By virtue of cleavable linker, the recombinant fusion polypeptides provided herein, such as the multimeric-fusion-proteins and redox fusion polypeptides, can be isolated and purified by introducing an enzyme or chemical that cleaves said multimeric-fusion-protein and/or redox fusion polypeptide from said oil body,
20 thereby obtaining and/or isolating the desired protein. It is contemplated herein that the use of cleavable linker sequence downstream of the non-proteolytic linker/spacer sequence will improve the yield of protein recovery when isolating or purifying proteins using the methods provided herein.

The nucleic acid sequences encoding the first or second recombinant
25 polypeptide may be altered to improve expression levels for example, by optimizing the nucleic acids sequence in accordance with the preferred codon usage for the particular cell type which is selected for expression of the first and second recombinant polypeptide, or by altering of motifs known to destabilize mRNAs (see for example: PCT Patent Application 97/02352). Comparison of
30 the codon usage of the first and second recombinant polypeptide with codon usage of the host will enable the identification of codons that may be changed. For example, typically plant evolution has tended towards a preference for CG

-28-

rich nucleotide sequences while bacterial evolution has resulted in bias towards AT rich nucleotide sequences. By modifying the nucleic acid sequences to incorporate nucleic acid sequences preferred by the host cell, expression may be optimized. Construction of synthetic genes by altering codon usage is described
5 in for example PCT patent Application 93/07278. The first and second recombinant polypeptide can be altered using for example targeted mutagenesis, random mutagenesis (Shiraishi et al. (1998) Arch. Biochem. Biophys. 358: 104-115; Galkin et al. (1997) Protein Eng. 10: 687-690; Carugo et al. (1997) Proteins 28: 10-28; Hurley et al. (1996) Biochemistry 35: 5670-5678), gene
10 shuffling, and/or by the addition of organic solvent (Holmberg et al. (1999) Protein Eng. 12: 851-856). Any polypeptide spacers that are used in accordance with the methods and products provided herein may be altered in similar ways.

In particular embodiments provided herein, the recombinant polypeptides
15 or thioredoxin-related proteins capable of forming a multimeric-protein-complex are capable of forming a heteromultimeric-protein-complex. Examples of heteromultimeric-protein-complexes that contain polypeptide chains that repeatedly interact, either to activate, inactivate, oxidize, reduce, stabilize, etc., with one another, that can be produced in association with oil bodies using the
20 methods provided herein include those set forth in Figure 5. Accordingly, exemplary proteins for use in the heteromultimeric-protein-complexes and nucleic acid constructs encoding such, provided herein include, among others described herein, those set forth in Figure 5.

Other polypeptide regions that can be used in the first and/or second
25 recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins, provided herein include, among other, those immunoglobulin regions set forth in Table 1.

TABLE 1 - ANTIBODY HETERODIMERS

	<u>Class or molecule</u>	<u>Subunits</u>
	Fab	Variable region and first constant region of heavy chain and complete light chain
5	Fv	Variable regions of heavy and light antibody chains
	IgA	heavy chains, light chains and J (joining) chain
	IgG, IgD, IgE	heavy and light chains
	IgM	heavy chains, light chains and J (joining) chain
10	Antibody chain(s) and a toxin	Antibody chain(s) and a toxin
	Autoantigens, allergens and transplant antigens with an adjuvant or tolerogen	Autoantigens, allergens and transplant antigens with an adjuvant or tolerogen
15	Chimeras using antibody Fc domain	Receptor subunits fused to the constant region of antibody heavy chains

As set forth above, in one embodiment, exemplary heteromultimeric-protein-complexes and exemplary heteromultimeric-fusion-proteins provided herein comprise redox proteins, such as the thioredoxins and thioredoxin-reductases and immunoglobulins.

20

Oil-body-targeting-proteins

The nucleic acid sequence encoding the oil-body-targeting-protein that may be used in the methods and compositions provided herein may be any nucleic acid sequence encoding an oil-body-targeting-protein, protein fragment or peptide capable of association with first recombinant polypeptide, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and/or a first and/or second thioredoxin-related protein and the oil bodies. The nucleic acid sequence encoding the oil body targeting peptide may be synthesized or obtained from any biological source.

For example, in one embodiment the oil-body-targeting-protein is an immunoglobulin or an immunoglobulin derived molecule, for example, a bispecific single chain antibody. The generation of single chain antibodies and bi-specific single chain antibodies is known to the art (see, e.g., US Patents US 5,763,733,

-30-

US5,767,260 and US5,260,203). Nucleic acid sequences encoding single chain antibodies functioning as oil-body-targeting-proteins may be prepared from hybridoma cell lines expressing monoclonal antibodies raised against an oleosin as described by Alting-Mees et al (2000) IBC's Annual International Conference on Antibody Engineering, Poster #1. In order to attain specificity for the first recombinant polypeptide a nucleic acid sequence encoding a second single chain antibody prepared from a monoclonal raised against the first recombinant polypeptide may be prepared and linked to the anti-oleosin single chain antibody. In this embodiment the oil body associates with the first recombinant polypeptide through non-covalent interactions of the oil-body-targeting-protein with the first recombinant polypeptide and the oil body. Alternatively the first recombinant polypeptide may be prepared as a fusion protein with an oil-body-targeting-protein. For example, a nucleic acid sequence encoding a single chain antibody raised against an oleosin may be fused to a nucleic acid sequence encoding the first recombinant polypeptide

Non-immunoglobulin-based oil-body-targeting-proteins capable of association with the first recombinant polypeptide may be discovered and prepared using for example phage display techniques (Pharmacia Biotech Catalogue Number 27-9401-011 Recombinant Phage Antibody System Expression Kit).

Oil-body-targeting-proteins may also be chemically modified. For example, oleosins may be modified by changing chemical modification of the lysine residues using chemical agents such as biotinyl-N-hydroxysuccinimide ester resulting in a process referred to as biotinylation. Conveniently this is accomplished by *in vitro* biotinylation of the oil bodies. *In vivo* biotinylation may be accomplished using the biotinylation domain peptide from the biotin carboxy carrier protein of *E. coli* acetyl-CoA carboxylase (Smith et al. (1998) Nucl. Acids. Res. 26: 1414-1420). Avidin or streptavidin may subsequently be used to accomplish association of the redox protein with the oil body.

In a particular embodiment the oil-body-targeting-protein is an oil-body-protein such as for example an oleosin or a caleosin or a sufficient portion derived thereof capable of targeting to an oil body. Nucleic acid sequences

-31-

encoding oleosins are known to the art. These include for example the *Arabidopsis* oleosin (van Rooijen et al (1991) Plant Mol. Bio. 18:1177-1179); the maize oleosin (Qu and Huang (1990) J. Biol. Chem. Vol. 265 4:2238-2243); rapeseed oleosin (Lee and Huang (1991) Plant Physiol. 96:1395-1397); and the
5 carrot oleosin (Hatzopoulos et al (1990) Plant Cell Vol. 2, 457-467.). Caleosin nucleic acid sequences are also known to the art (Naested et al (2000) Plant Mol Biol. 44(4):463-476; Chen et al (1999) Plant Cell Physiol. 40(10):1079-1086). Animal cell derived oil body proteins that may be used herein include adipophilin (Brasaemle et al, (1997) J. Lipid Res., 38: 2249-2263; Heid et al. (1998) Cell
10 Tissue Research 294: 309-321), perilipin (Blanchette-Mackie et al. (1995), J. Lipid Res. 36: 1211-1226; Servetnick et al. (1995) J. Biol. Chem. 270: 16970-16973), apolipoproteins such as apo A-I, A-II, A-IV, C-I, C-II, CIII (Segrest et al. (1990), Proteins 8:103-117) and apoB (Chatterton et al. (1995) J. Lipid Res. 36: 2027-2037; Davis, RA in: Vance DE, Vance J. editors. Lipoprotein structure and
15 secretion. The Netherlands, Elsevier, 191: 403-426.

In one embodiment, the first recombinant polypeptide is fused to an oil-body-protein. The methodology is further described in US patent 5,650,554, which is incorporated herein by reference in its entirety. The first recombinant polypeptide may be fused to the N-terminus as well as to the C-terminus of the
20 oil-body-protein (as described in: Moloney and van Rooijen (1996) INFORM 7:107-113) and fragments of the oil-body-protein such as for example the central domain of an oleosin molecule, or modified versions of the oil-body-protein may be used. In this embodiment, the second recombinant polypeptide is expressed intracellularly and then intracellularly associates with the first
25 recombinant polypeptide to form the multimeric-protein-complex in the cell. Oil bodies comprising the multimeric-protein-complex are then conveniently isolated from the cells.

In a further embodiment both the first and second recombinant polypeptide are separately fused to an oil-body-protein. In this embodiment
30 nucleic acid sequences encoding the first and second polypeptides may be prepared separately and introduced in separate cell lines or they may be introduced in the same cell lines. Where the nucleic acid sequences are

-32-

introduced in the same cell line, these nucleic acid sequence may be prepared using two separate expression vectors, or they may be prepared using a single vector comprising nucleic acid sequences encoding both the first polypeptide fused to an oil body protein and the second polypeptide fused to an oil-body-
5 protein. Where separate cell lines are used subsequent mating of the offspring (e.g., mating of plants) is used to prepare a generation of cells comprising oil bodies which comprise both the first and second recombinant polypeptide fused to an oil-body-protein.

In further alternate embodiment, the first and second recombinant
10 polypeptide are fused to form a multimeric-fusion-protein comprising the multimeric-protein-complex. In such an embodiment, the first and second polypeptide is associated with the oil body through an oil-body-targeting-protein capable of associating with both the fusion protein and with the oil body. In a particular embodiment, the fusion protein comprising the multimeric-protein-
15 complex is fused to an oil-body-protein, for example, an oleosin or caleosin.

In embodiments provided herein in which the multimeric-protein-complex is an immunoglobulin (e.g., a multimeric-immunoglobulin-complex), a particularly preferred oil body targeting protein is an oleosin or caleosin associated with an immunoglobulin binding protein, such as for example protein A (US Patent
20 5,151,350), protein L (US Patent 5,965,390) and protein G (US Patent 4,954,618), or active fragments of such immunoglobulin binding proteins.

New oil-body-proteins may be discovered for example by preparing oil bodies (described in further detail below) and identifying proteins in these preparations using for example SDS gel electrophoresis. Polyclonal antibodies
25 may be raised against these proteins and used to screen cDNA libraries in order to identify nucleic acid sequences encoding oil-body-proteins. The methodologies are familiar to the skilled artisan (Huynh et al. (1985) in DNA Cloning Vol. 1. a Practical Approach ed. DM Glover, IRL Press, pp 49-78). New oil-body-proteins may further be discovered using known nucleic acid sequences
30 encoding oil-body-proteins (e.g. the *Arabidopsis*, rapeseed, carrot and corn nucleic acid sequences) to probe for example cDNA and genomic libraries for the presence of nucleic acid sequences encoding oil-body-proteins.

-33-

In one embodiment, the first and second polypeptide are a first and second redox protein. Accordingly, one embodiment provided herein relates to novel and improved methods for the production of redox proteins. It has unexpectedly been found that a redox protein when prepared as a fusion protein with a second redox protein is fully enzymatically active when produced in association with an oil body. In contrast, when the redox protein is prepared without the second redox protein it has reduced enzymatic activity. In one embodiment, the first redox protein is at least 5 times more active when produced as a redox fusion polypeptide relative to production as a non-fusion polypeptide.

Accordingly, provided herein are methods for producing an oil body associated with a heteromultimeric redox protein complex, said method comprising:

(a) producing in a cell comprising oil bodies, a first redox protein and a second redox protein wherein said first redox protein is capable of interacting with said second redox protein, preferably in the cell, to form said heteromultimeric redox protein complex; and

(b) associating said heteromultimeric redox protein complex with an oil body through an oil-body-targeting-protein capable of associating with said oil bodies and said heteromultimeric redox protein complex.

In a particular embodiment the first and second redox protein are prepared as a fusion protein to form a redox fusion polypeptide. Accordingly, provided herein are methods for preparing an enzymatically active redox protein associated with oil bodies comprising:

a) producing in a cell a redox fusion polypeptide comprising a first redox protein linked to a second redox protein;

b) associating said redox fusion polypeptide with oil bodies through an oil-body-targeting-protein capable of associating with said redox fusion polypeptide and said oil bodies; and

c) isolating said oil bodies associated with said redox fusion polypeptide.

The oil bodies in association with the redox protein may be used to prepare a variety of useful emulsions.

-34-

As used herein the phrase "redox proteins" or grammatical variations thereof, refers to any protein or active protein fragment capable of participating in electron transport. For example, redox proteins are capable of catalyzing the transfer of an electron from an electron donor (also frequently referred to as the
5 reducing agent) to an electron acceptor (also frequently referred to as the oxidizing agent). In the process of electron transfer, the reducing agent (electron donor) is oxidized and the oxidizing agent (electron acceptor) is reduced. Exemplary redox proteins for use herein include iron-sulfur proteins, cytochromes, redox active thiol proteins and redox-active flavoproteins. To
10 carry out their function as conduits for electrons, redox proteins, such as thioredoxin and thioredoxin-reductase for example, are known to function by interacting or associating with one another in multimeric-protein-complexes (e.g., heteromultimeric-protein-complexes).

The term "redox fusion polypeptide" as used herein refers to any fusion
15 polypeptide comprising a first redox protein linked to a second redox protein (e.g., an in-frame translational fusion). The redox proteins that may be used with the methods and compositions provided herein may be any redox protein. In one embodiment the first and second redox proteins are a pair of redox proteins that would normally occur together from the same source, in nature. In
20 a particular embodiment, the first redox protein is a thioredoxin and the second redox protein is a thioredoxin-reductase.

The redox fusion polypeptide may be produced in any cell comprising oil bodies, including any animal cell, plant cell, algae cell, fungal cell or bacterial cell. In certain embodiments the redox fusion polypeptide is produced in a plant
25 cell and in particular embodiments the redox fusion polypeptide is produced in the seed cells of a seed plant.

In particular embodiments the oil-body-targeting-protein that is used is an oil-body-protein. In embodiments of the present invention in which an oil-body-protein is used, the first and second redox protein are preferably covalently fused
30 to the oil-body-protein. Accordingly, provided herein are methods for the preparation of a redox protein in association with an oil body comprising:

- a) introducing into a cell a chimeric nucleic acid sequence comprising:

-35-

- 1) a first nucleic acid sequence capable of regulating transcription in said cell operatively linked to;
- 2) a second nucleic acid sequence encoding a recombinant fusion polypeptide comprising (i) a first nucleic acid sequence encoding a sufficient portion of an oil-body-protein to provide targeting of said recombinant fusion polypeptide to an oil body linked in reading frame to (ii) a second nucleic acid sequence encoding a redox fusion polypeptide comprising a first redox protein linked to a second redox protein operatively linked to;
- 3) a third nucleic acid sequence capable of terminating transcription in said cell;
- b) growing said cell under conditions to permit expression of said redox fusion polypeptide in a progeny cell comprising oil bodies; and
- c) isolating said oil bodies comprising said redox fusion polypeptide from said progeny cell.

Redox Proteins

In accordance with various methods and compositions provided herein, any nucleic acid sequence encoding a redox protein may be used. The nucleic acid sequence encoding the first and/or second redox protein may be obtained from any biological source or may be prepared synthetically. In general, nucleic acid sequences encoding redox proteins are well known in the art and readily available. See, for example: Cristiano et al. (1993) *Genomics* 17: (2) 348-354, Doyama et al. (1998) *Plant Sci.* 137: 53-62, Hoeoeg et al. (1984) *Biosci. Rep.* 4: 917-923; as well as the Swiss Protein sequences set forth in Table 5. Known nucleic acid sequences encoding redox proteins may be used to design and construct nucleic acid sequence based probes in order to uncover and identify previously undiscovered nucleic acid sequences encoding redox proteins, for example by screening cDNA or genomic libraries. Thus, additional nucleic acid sequences may be discovered and used in accordance with the present invention.

-36-

The nucleic acid sequence encoding the first and/or second redox protein may be obtained from separate sources or may be obtained from the same source. In general however, the nucleic acid sequence encoding a redox-fusion polypeptide comprises nucleic acid sequences encoding a first and a second
5 redox protein obtained from the same or a similar biological source. In certain embodiments provided herein, wherein the nucleic acid sequence encoding the first and second redox protein is obtained from the same source, the nucleic acid sequence encoding the first redox protein and second redox protein may be naturally fused. In accordance with a particular embodiment, the nucleic acid
10 sequences encoding the first and second redox protein are preferably obtained from a plant source.

As set forth above, a polypeptide spacer or linker of variable length may separate the first and second redox proteins from each other and/or from the oil-body-targeting-protein; and additional redox proteins (e.g., one or more) may be
15 fused to the first and/or second redox protein.

The nucleic acid sequences encoding the redox proteins may be altered to improve expression levels for example by optimizing the nucleic acids sequence in accordance with the preferred codon usage for the particular cell type which is selected for expression of the redox proteins, or by altering of motifs known to
20 destabilize mRNAs (see for example: PCT Patent Application 97/02352).

Comparison of the codon usage of the redox protein with codon usage of the host will enable the identification of codons that may be changed. For example, typically plant evolution has tended towards a preference for CG rich nucleotide sequences while bacterial evolution has resulted in bias towards AT rich
25 nucleotide sequences. By modifying the nucleic acid sequences to incorporate nucleic acid sequences preferred by the host cell, expression may be optimized. Construction of synthetic genes by altering codon usage is described in for example PCT patent Application 93/07278. The redox proteins may be altered using for example, targeted mutagenesis, random mutagenesis (Shiraishi et al.
30 (1998) Arch. Biochem. Biophys. 358: 104-115; Galkin et al. (1997) Protein Eng. 10: 687-690; Carugo et al. (1997) Proteins 28: 10-28; Hurley et al. (1996) Biochemistry 35: 5670-5678) (and/or by the addition of organic solvent

-37-

(Holmberg et al. (1999) Protein Eng. 12: 851-856). The polypeptide spacer between the first and second redox protein may be altered in similar ways.

The first and second redox protein may be selected by developing a two-dimensional matrix and determining which combination of first and second redox protein is most effective in electron transport using for example, a colorimetric reduction assay (Johnson et al (1984) J. of Bact. Vol. 158 3:1061-1069, Luthman et al (1982) Biochemistry Vol 21 26:6628-2233). Combinations of thioredoxin and thioredoxin-reductase may be tested by determining the reduction of wheat storage proteins and milk storage protein beta-lactoglobulin in vitro (Del Val et al. (1999) J. Allerg. Clin. Immunol. 103: 690-697). Using the same strategy polypeptide spacers between the first and second redox proteins may be evaluated for their efficiency.

First and second redox proteins that may be used herein include without limitation any first redox protein and second redox protein selected from the group of redox proteins consisting of cytochromes, such as cytochrome a, cytochrome b and cytochrome c; porphyrin containing proteins, for example hemoglobin; iron-sulfur proteins, such as ferredoxin; flavoproteins such as thioredoxin-reductase, NADH dehydrogenase, succinate dehydrogenase, dihydrolipoyl dehydrogenase, acyl-CoA dehydrogenase, D-amino acid oxidase, xanthine oxidase, orotate reductase and aldehyde oxidase; pyridine-linked dehydrogenases, for example, lactate dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase, malate dehydrogenase, and beta-hydroxy-butarate dehydrogenase; and redox active thiol containing proteins such as thioredoxin.

In particular embodiments, the redox proteins provided herein are thioredoxin and its reductant thioredoxin-reductase (which are jointly also referred to herein as "thioredoxin-related" protein(s)). As used herein, the term "thioredoxin" refers to relatively small proteins (typically approximately 12 kDa) that belong to the family of thioltransferases which catalyze oxido-reductions via the formation or hydrolysis of disulfide bonds and are widely, if not universally, distributed throughout the animal plant and bacterial kingdom. The reduced form of thioredoxin is an excellent catalyst for the reduction of even the most intractable disulfide bonds. In order to reduce the oxidized thioredoxin, two

-38-

cellular reductants provide the reduction equivalents: reduced ferredoxin and NADPH. These reduction equivalents are supplied to thioredoxin via interaction or association with different thioredoxin-reductases including the NADPH thioredoxin-reductase and ferredoxin thioredoxin-reductase. The supply of these

5 reduction equivalents requires the formation of a heteromultimeric-protein-complex comprising thioredoxin and thioredoxin-reductase. Ferredoxin thioredoxin-reductase is involved in the reduction of plant thioredoxins designated as Trxf and Trxm, both of which are involved in the regulation of photosynthetic processes in the chloroplast. The NADPH/thioredoxin active in

10 plant seeds is designated Trxh (also referred to herein as thioredoxin h-type) and is capable of the reduction of a wide range of proteins thereby functioning as an important cellular redox buffer. Generally, only one kind of thioredoxin, which analogous to the plant Trxh type, is found in bacterial or animal cells. The h-type thioredoxins are capable of being reduced by NADPH and NADPH-

15 thioredoxin reductase.

Exemplary thioredoxins are further characterized as a protein having a core of 5 beta-sheets surrounded by 4 to 6 alpha helixes. Exemplary thioredoxins are further characterized by having an active site containing the consensus amino acid sequence:

20 X C Y Y C Z,

wherein Y is any amino acid, such as hydrophobic or non-polar amino acids, wherein X can be any of the 20 amino acids, preferably a hydrophobic amino acid, such as a tryptophan, and Z can be any amino acid, preferably polar amino acids.

25 In certain embodiments, the thioredoxins for use herein comprise an active site having the amino acid sequence X C G P C Z.

When the cysteines in the active site of thioredoxin or thioredoxin-like proteins are oxidized, they form an intramolecular disulfide bond. In the reduced state, the same active sites are capable of participating in redox reactions through the

30 reversible oxidation of its active site dithiol, to a disulfide and catalyzes dithiol-disulfide exchange reactions.

Exemplary thioredoxins are well-known in the art and can be obtained from several organisms including *Arabidopsis thaliana* (Riveira Madrid et al. (1995) Proc. Natl. Acad. Sci. 92: 5620-5624), wheat (Gautier et al. (1998) Eur. J. Biochem. 252: 314-324); *Escherichia coli* (Hoeoeg et al (1984) Biosci. Rep. 4: 917-923) and thermophylic microorganisms such as *Methanococcus jannaschii* and *Archaeoglobus fulgidus* (PCT Patent Application 00/36126). Thioredoxins have also been recombinantly expressed in several host systems including bacteria (Gautier et al. (1998) Eur J. Biochem. 252: 314-324) and plants (PCT Patent Application WO 00/58453) Commercial preparations of *E. coli* sourced Thioredoxins are readily available from for example: Sigma Cat No. T 0910 Thioredoxin (*E. coli*, recombinant; expressed in *E. coli*).

Exemplary nucleic acid sequences encoding thioredoxin polypeptides for use herein are readily available from a variety of diverse biological sources including *E. coli* (Hoeoeg et al. (1984) Biosci. Rep.: 4 917-923); *Methanococcus jannaschii* and *Archaeoglobus fulgidus* (PCT Patent Application 00/36126); *Arabidopsis thaliana* (Rivera-Madrid (1995) Proc. Natl. Acad. Sci. 92: 5620-5624); wheat (Gautier et al (1998) Eur. J. Biochem. 252(2): 314-324); tobacco (Marty et al. (1991) Plant Mol. Biol. 17: 143-148); barley (PCT Patent Application 00/58352); rice (Ishiwatari et al. (1995) Planta 195: 456-463); soybean (Shi et al. (1996) Plant Mol. Biol. 32: 653-662); rapeseed (Bower et al. Plant Cell 8: 1641-1650) and calf (Terashima et al. (1999) DNA Seq. 10(3): 203-205); and the like. In yet other embodiments, exemplary nucleic acids for use herein include those encoding the thioredoxin and thioredoxin-like polypeptide chains set forth as SEQ ID NOs:38, 42, 46 and 50; and those encoding the thioredoxin and thioredoxin-like polypeptide chains set forth in Table 5 as SEQ ID NOs:52-194. The respective nucleic acid sequences encoding the amino acids set forth in SEQ ID NOs:52-194 can be readily identified via the Swiss Protein identifier (accession) numbers provided in Table 5 (in parenthesis).

As used herein, the term "thioredoxin-reductase" refers to a protein that complexes with a flavin, such as FAD. The flavin compound serves as an electron donor for the thioredoxin-reductase protein active site. Thioredoxin

-40-

reductases have a redox active, disulfide bond site capable of reducing thioredoxin. The active site of thioredoxin-reductase contains 2 cysteines. The type of amino acids surrounding the 2 cysteine residues forming the active site can vary as hydrophobic, non-polar or polar. An exemplary thioredoxin-

5 reductase is NADPH-thioredoxin-reductase (NTR), which is a cytosolic homodimeric enzyme comprising typically 300-500 amino acids. Crystal structures of both *E. coli* and plant thioredoxin-reductase have been obtained (Waksman et al. (1994) J. Mol. Biol. 236: 800-816; Dai et al. (1996) J. Mol. Biol. 264:1044-1057). NADPH-thioredoxin-reductases have been expressed in

10 heterologous hosts, for example the *Arabidopsis* NADPH-thioredoxin-reductase has been expressed in *E. coli* (Jacquot et al. (1994) J. Mol. Biol. 235: 1357-1363) and wheat (PCT Patent Application 00/58453).

Exemplary nucleic acid sequences encoding thioredoxin-reductase proteins can readily be obtained from a variety of sources, such as from the

15 sequence set forth in Table 5 and the Sequence Listing provide herein, from *Arabidopsis* (Riveira Madrid et al. (1995) Proc. Natl. Acad. Sci. USA 92: 5620-5624), *E. coli* (Russel et al. (1988) J. Biol. Chem. 263: 9015-9019); barley (PCT Patent Application 00/58352 and wheat (Gautier et al., (1998) Eur. J. Biochem. 252: 314-324); and the like. In yet other embodiments, exemplary nucleic acids

20 for use herein include those encoding the thioredoxin-reductase polypeptide chains set forth as SEQ ID NOs:8, 9, 10, 40, 44, 48 and 50; and those encoding the thioredoxin-reductase polypeptide chains set forth in Table 5 as SEQ ID NOs:195-313. The respective nucleic acid sequences encoding the amino acids set forth in SEQ ID NOs:195-313 can be readily identified via the

25 Swiss Protein identifier (accession) numbers provided in Table 5 (in parenthesis).

Also contemplated for use in the methods and compositions provided herein are nucleic acid and amino acid homologs that are "substantially homologous" to the thioredoxin and thioredoxin-reductase nucleic and amino acids set forth herein, which includes thioredoxin and thioredoxin-reductase

30 polypeptides encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides encoding the thioredoxin and thioredoxin-reductase nucleic and amino acids set

-41-

forth herein (e.g., in the Examples, Sequence Listing and/or Table 5). As used herein, a DNA or nucleic acid homolog refers to a nucleic acid that includes a preselected conserved nucleotide sequence, such as a sequence encoding a therapeutic polypeptide. By the term "substantially homologous" is meant
5 having at least 80%, preferably at least 90%, most preferably at least 95% homology therewith or a less percentage of homology or identity and conserved biological activity or function.

The terms "homology" and "identity" are often used interchangeably. In this regard, percent homology or identity may be determined, for example, by
10 comparing sequence information using a GAP computer program. The GAP program utilizes the alignment method of Needleman and Wunsch (*J. Mol. Biol.* 48:443 (1970), as revised by Smith and Waterman (*Adv. Appl. Math.* 2:482 (1981)). Briefly, the GAP program defines similarity as the number of aligned symbols (i.e., nucleotides or amino acids) which are similar, divided by the total
15 number of symbols in the shorter of the two sequences. The preferred default parameters for the GAP program may include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) and the weighted comparison matrix of Gribskov and Burgess, *Nucl. Acids Res.* 14:6745 (1986), as described by Schwartz and Dayhoff, eds., *ATLAS OF PROTEIN SEQUENCE*
20 *AND STRUCTURE*, National Biomedical Research Foundation, pp. 353-358 (1979); (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps.

By sequence identity, the number of conserved amino acids are determined by standard alignment algorithms programs, and are used with
25 default gap penalties established by each supplier. Substantially homologous nucleic acid molecules would hybridize typically at moderate stringency or at high stringency all along the length of the nucleic acid of interest. Preferably the two molecules will hybridize under conditions of high stringency. Also contemplated are nucleic acid molecules that contain degenerate codons in place
30 of codons in the hybridizing nucleic acid molecule.

Whether any two nucleic acid molecules have nucleotide sequences that are at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% "identical" can be

-42-

determined using known computer algorithms such as the "FAST A" program, using for example, the default parameters as in Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444 (1988). Alternatively the BLAST function of the National Center for Biotechnology Information database may be used to
5 determine relative sequence identity.

In general, sequences are aligned so that the highest order match is obtained. "Identity" *per se* has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics*
10 *and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991).
15 While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans (Carillo, H. & Lipton, D., *SIAM J Applied Math* 48:1073 (1988)). Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge
20 Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H. & Lipton, D., *SIAM J Applied Math* 48:1073 (1988). Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux, J.,
25 *et al.*, *Nucleic Acids Research* 12(II):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F., *et al.*, *J Molec Biol* 215:403 (1990)).

Therefore, as used herein, the term "identity" represents a comparison between a test and a reference polypeptide or polynucleotide. For example, a test polypeptide may be defined as any polypeptide that is 90% or more
30 identical to a reference polypeptide.

As used herein, the term at least "90% identical to" refers to percent identities from 90 to 99.99 relative to the reference polypeptides. Identity at a

-43-

level of 90% or more is indicative of the fact that, assuming for exemplification purposes a test and reference polynucleotide length of 100 amino acids are compared. No more than 10% (i.e., 10 out of 100) amino acids in the test polypeptide differs from that of the reference polypeptides. Similar comparisons
5 may be made between a test and reference polynucleotides. Such differences may be represented as point mutations randomly distributed over the entire length of an amino acid sequence or they may be clustered in one or more locations of varying length up to the maximum allowable, e.g. 10/100 amino acid difference (approximately 90% identity). Differences are defined as nucleic
10 acid or amino acid substitutions, or deletions.

As used herein: stringency of hybridization in determining percentage mismatch is as follows:

- 1) high stringency: 0.1 x SSPE, 0.1% SDS, 65°C
- 2) medium stringency: 0.2 x SSPE, 0.1% SDS, 50°C
- 15 3) low stringency: 1.0 x SSPE, 0.1% SDS, 50°C

Those of skill in this art know that the washing step selects for stable hybrids and also know the ingredients of SSPE (see, e.g., Sambrook, E.F. Fritsch, T. Maniatis, in: Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press (1989), vol. 3, p. B.13, see, also, numerous catalogs
20 that describe commonly used laboratory solutions). SSPE is pH 7.4 phosphate-buffered 0.18 M NaCl. Further, those of skill in the art recognize that the stability of hybrids is determined by T_m , which is a function of the sodium ion concentration and temperature ($T_m = 81.5^\circ \text{C} - 16.6(\log_{10}[\text{Na}^+]) + 0.41(\% \text{G} + \text{C}) - 600/l$), so that the only parameters in the wash conditions critical to hybrid
25 stability are sodium ion concentration in the SSPE (or SSC) and temperature.

It is understood that equivalent stringencies may be achieved using alternative buffers, salts and temperatures. By way of example and not limitation, procedures using conditions of low stringency are as follows (see also Shilo and Weinberg, *Proc. Natl. Acad. Sci. USA*, 78:6789-6792 (1981)): Filters
30 containing DNA are pretreated for 6 hours at 40°C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.1% PVP, 0.1%

-44-

Ficoll, 1% BSA, and 500 μ g/ml denatured salmon sperm DNA (10X SSC is 1.5 M sodium chloride, and 0.15 M sodium citrate, adjusted to a pH of 7).

In a particular embodiment, a heteromultimeric-protein-complex is produced as a fusion polypeptide between the first and second redox protein, wherein the first redox protein is thioredoxin and the second redox protein is a thioredoxin-reductase. In one embodiment, the second recombinant polypeptide, e.g., the thioredoxin-reductase is positioned N-terminal relative to the first recombinant polypeptide, e.g., the thioredoxin. Accordingly, any protein which is classified as thioredoxin, such as the thioredoxin component of the NADPH thioredoxin system and the thioredoxin present in the ferredoxin/thioredoxin system also known as TRx and TRm may be used in combination with any thioredoxin-reductase such as the NADPH thioredoxin-reductase and the ferredoxin-thioredoxin-reductase and any other proteins having the capability of reducing thioredoxin. In particular embodiments the thioredoxin and thioredoxin-reductase are plant derived.

In an alternate embodiment, the naturally occurring nucleic acid sequence encoding the thioredoxin/thioredoxin-reductase protein fusion obtainable from *Mycobacterium leprae* (Wieles et al. (1995) J. Biol. Chem. 27:25604-25606) is used, as set forth in the Examples herein.

20 Immunoglobulins

In another embodiment of the present invention, the multimeric-protein-complexes are immunoglobulins. As used herein "immunoglobulin-polypeptide-chain" refers to a first polypeptide that is capable of associating with a second polypeptide to form an immunologically active (i.e. capable of antigen binding) multimeric-protein-complex. The types of immunoglobulins and immunoglobulin-polypeptide-chains contemplated for use herein include the immunologically active (i.e. antigen binding) portions of a light and heavy chain. Exemplary immunoglobulins and immunoglobulin-polypeptide-chains for use herein include substantially intact immunoglobulins, including any IgG, IgA, IgD, IgE and IgM, as well as any portion of an immunoglobulin, including those portions well-known as Fab fragments, Fab' fragments, F(ab')₂ fragments and Fv fragments.

-45-

In this embodiment, the first recombinant polypeptide may be any immunoglobulin heavy chain, including any IgG, IgA, IgD, IgE or IgM heavy chain, and the second recombinant polypeptide may be a kappa or lambda immunoglobulin light chain. Accordingly, provided herein are methods of

5 producing an immunoglobulin, said method comprising: (a) producing in a cell comprising oil bodies, a first immunoglobulin-polypeptide-chain and a second immunoglobulin-polypeptide-chain wherein said first immunoglobulin-polypeptide-chain is capable of associating with said second immunoglobulin-polypeptide-chain to form said immunoglobulin; and (b) associating said immunoglobulin with

10 an oil body through an oil-body-targeting-protein capable of associating with said oil bodies and said first immunoglobulin-polypeptide-chain.

As set forth herein, the multimeric immunoglobulin is associated with an oil body through an oil-body-targeting-protein. In particular embodiments, the oil-body-targeting-protein may be a fusion polypeptide comprising an

15 oil-body-protein and an immunoglobulin binding protein, such as for example protein A, protein L, and protein G.

In yet another embodiment involving immunoglobulins, the first and second recombinant polypeptides (immunoglobulins) are separately fused to an oil body protein, for example an oleosin or caleosin. For example,

20 a) the first recombinant polypeptide may be an immunoglobulin heavy chain, including any IgG, IgA, IgD, IgE or IgM heavy chain, and the second recombinant polypeptide may be a kappa or lambda immunoglobulin light chain; or

b) the first recombinant polypeptide may be the variable and first

25 constant domain from an immunoglobulin heavy chain and the second recombinant polypeptide may be a kappa or lambda immunoglobulin light chain; or

c) the first recombinant polypeptide may be the variable domain from an immunoglobulin heavy chain and the second recombinant polypeptide may be

30 the variable domain from a kappa or lambda immunoglobulin light chain.

In certain embodiments, the fusion polypeptides are designed or selected to allow the heteromultimeric-protein-complex formation between

-46-

immunoglobulin light and heavy chain sequences on the oil bodies within the cell comprising oil bodies.

Preparation of expression vectors comprising oil-body-targeting-proteins and the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins.

In accordance with the present invention, the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins; and the oil-body-targeting-protein are conveniently produced in a cell. In order to produce the recombinant polypeptides or multimeric-protein-complexes, a nucleic acid sequence encoding either the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins; and/or the oil-body-targeting-protein are incorporated in a recombinant expression vector. Accordingly, provided herein are recombinant expression vectors comprising the chimeric nucleic acids provided herein suitable for expression of the oil-body-targeting-protein and the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins, suitable for the selected cell. The term "suitable for expression in the selected cell" means that the recombinant expression vector contains all nucleic acid sequences required to ensure expression in the selected cell.

Accordingly, the recombinant expression vectors further contain regulatory nucleic acid sequences selected on the basis of the cell which is used

-47-

for expression and ensuring initiation and termination of transcription operatively linked to the nucleic acid sequence encoding the recombinant polypeptide or multimeric-protein-complex and/or the oil-body-targeting-protein. Regulatory nucleic acid sequences include promoters, enhancers, silencing elements, ribosome binding sites, Shine-Dalgarno sequences, introns and other expression elements. "Operatively linked" is intended to mean that the nucleic acid sequences comprising the regulatory regions linked to the nucleic acid sequences encoding the recombinant polypeptide or multimeric-protein-complex and/or the oil-body-targeting-protein allow expression in the cell. A typical nucleic acid construct comprises in the 5' to 3' direction a promoter region capable of directing expression, a coding region comprising the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins; and/or an oil-body-targeting-protein and a termination region functional in the selected cell.

The selection of regulatory sequences will depend on the organism and the cell type in which the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins; and/or the oil-body-targeting-protein is expressed, and may influence the expression levels of the polypeptide. Regulatory sequences are art-recognized and selected to direct expression of the oil-body-targeting-protein and the recombinant polypeptides or multimeric-protein-complexes in the cell.

Promoters that may be used in bacterial cells include the lac promoter (Blackman et al. (1978) Cell: 13: 65-71), the trp promoter (Masuda et al. (1996) Protein Eng: 9: 101-106) and the T7 promoters (Studier et al. (1986) J. Mol. Biol. 189: 113-130). Promoters functional in plant cells that may be used herein include constitutive promoters such as the 35S CaMV promoter (Rothstein et al. (1987) Gene: 53: 153-161) the actin promoter (McElroy et al. (1990) Plant Cell

-48-

2: 163-171) and the ubiquitin promoter (European Patent Application 0 342 926). Other promoters are specific to certain tissues or organs (for example, roots, leaves, flowers or seeds) or cell types (for example, leaf epidermal cells, mesophyll cells or root cortex cells) and or to certain stages of plant development. Timing of expression may be controlled by selecting an inducible promoter, for example the PR-a promoter described in US Patent 5,614,395. Selection of the promoter therefore depends on the desired location and timing of the accumulation of the desired polypeptide. In a particular embodiment, the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins; and the oil-body-targeting-protein are expressed in a seed cell and seed specific promoters are utilized. Seed specific promoters that may be used herein include for example the phaseolin promoter (Sengupta-Gopalan et al. (1985) Proc. Natl. Acad. Sci. USA: 82: 3320-3324), and the *Arabidopsis* 18 kDa oleosin promoter (van Rooijen et al. (1992) Plant. Mol. Biol. 18: 1177-1179). New promoters useful in various plant cell types are constantly discovered. Numerous examples of plant promoters may be found in Ohamuro et al. (Biochem of Pl. (1989) 15: 1-82).

Genetic elements capable of enhancing expression of the polypeptide may be included in the expression vectors. In plant cells these include for example, the untranslated leader sequences from viruses such as the AMV leader sequence (Jobling and Gehrke (1987) Nature: 325: 622-625) and the intron associated with the maize ubiquitin promoter (See: US Patent 5,504,200).

Transcriptional terminators are generally art recognized and besides serving as a signal for transcription termination serve as a protective element serving to extend the mRNA half-life (Guarneros et al. (1982) Proc. Natl. Acad. Sci. USA: 79: 238-242). In nucleic acid sequences for the expression in plant cells, the transcriptional terminator typically is from about 200 nucleotide to about 1000 nucleotides in length. Terminator sequences that may be used herein include for example, the nopaline synthase termination region (Bevan et

-49-

al. (1983) Nucl. Acid. Res.: 11: 369-385), the phaseolin terminator (van der Geest et al. (1994) Plant J.: 6: 413-423), the terminator for the octopine synthase gene of *Agrobacterium tumefaciens* or other similarly functioning elements. Transcriptional terminators can be obtained as described by An
5 (1987) Methods in Enzym. 153: 292). The selection of the transcriptional terminator may have an effect on the rate of transcription.

Accordingly, provided herein are chimeric nucleic acid sequences encoding a first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins,
10 heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and/or thioredoxin-related proteins. In one embodiment, said nucleic acid comprises:

- (a) a first nucleic acid sequence encoding an oil-body-targeting-protein operatively linked in reading frame to;
- 15 (b) a second nucleic acid sequence encoding a first recombinant polypeptide, immunoglobulin-polypeptide-chain, or redox protein; linked in reading frame to;
- (c) a third nucleic acid sequence encoding a second recombinant polypeptide, immunoglobulin-polypeptide-chain or redox protein, wherein said
20 first and second recombinant polypeptides, immunoglobulin-polypeptide-chains or redox proteins are capable of forming a multimeric-protein-complex.

In another embodiment, provided herein is an expression vector comprising:

- 1) a first nucleic acid sequence capable of regulating transcription in said
25 cell operatively linked to;
- 2) a second nucleic acid sequence encoding a recombinant fusion polypeptide comprising (i) a nucleic acid sequence encoding a sufficient portion of an oil-body-protein to provide targeting of said recombinant fusion polypeptide to an oil body linked in reading frame to (ii) a nucleic acid sequence encoding a
30 multimeric-fusion-protein, such as a redox fusion polypeptide or immunoglobulin, comprising a first recombinant polypeptide, such as a redox protein or immunoglobulin-polypeptide-chain, linked to a second recombinant polypeptide,

-50-

such as a second redox protein or a second immunoglobulin-polypeptide-chain, operatively linked to;

3) a third nucleic acid sequence capable of terminating transcription in said cell.

- 5 The recombinant expression vector further may contain a marker gene. Marker genes that may be used in accordance with the present invention include all genes that allow the distinction of transformed cells from non-transformed cells including all selectable and screenable marker genes. A marker may be a resistance marker such as an antibiotic resistance marker against for example
- 10 kanamycin, ampicillin, G418, bleomycin hygromycin, chloramphenicol which allows selection of a trait by chemical means or a tolerance marker against for example a chemical agent such as the normally phytotoxic sugar mannose (Negrotto et al. (2000) Plant Cell Rep. 19: 798-803). In plant recombinant expression vectors herbicide resistance markers may conveniently be used for
- 15 example markers conferring resistance against glyphosate (US Patents 4,940,935 and 5,188,642) or phosphinothricin (White et al. (1990) Nucl. Acids Res. 18: 1062; Spencer et al. (1990) Theor. Appl. Genet. 79: 625-631). Resistance markers to a herbicide when linked in close proximity to the redox protein or oil-body-targeting-protein may be used to maintain selection pressure
- 20 on a population of plant cells or plants for those plants that have not lost the protein of interest. Screenable markers that may be employed to identify transformants through visual observation include beta-glucuronidase (GUS) (see US Patents US5,268,463 and US5,599,670) and green fluorescent protein (GFP) (Niedz et al. (1995) Plant Cell Rep.: 14: 403).
- 25 The recombinant expression vectors further may contain nucleic acid sequences encoding targeting signals ensuring targeting to a cell compartment or organelle. Suitable targeting signals that may be used herein include those that are capable of targeting polypeptides to the endomembrane system. Exemplary targeting signals that may be used herein include targeting signals capable of
- 30 directing the protein to the periplasm, the cytoplasm, the golgi apparatus, the apoplast (Sijmons et al., 1990, Bio/Technology, 8:217-221) the chloroplast (Comai et al. (1988) J. Biol. Chem. 263: 15104-15109), the mitochondrion, the

-51-

peroxisome (Unger et al. (1989) Plant Mol. Biol. 13: 411-418), the ER, the vacuole (Shinshi et al. (1990) Plant Mol. Biol. 14: 357-368 and the oil body. By the inclusion of the appropriate targeting sequences it is possible to direct the oil-body-targeting-protein or the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and/or thioredoxin-related proteins, to the desired organelle or cell compartment.

The recombinant expression vectors of the present invention may be prepared in accordance with methodologies well known to those of skill in the art of molecular biology (see for example: Sambrook et al. (1990) Molecular Cloning, 2nd ed. Cold Spring Harbor Press). The preparation of these constructs may involve techniques such as restriction digestion, ligation, gel electrophoresis, DNA sequencing and PCR. A wide variety of cloning vectors is available to perform the necessary cloning steps resulting in a recombinant expression vector ensuring expression of the polypeptide. Especially suitable for this purpose are vectors with a replication system that is functional in *Escherichia coli* such as pBR322, the PUC series of vectors, the M13mp series of vectors, pBluescript etc. Typically these vectors contain a marker allowing the selection of transformed cells for example by conferring antibiotic resistance. Nucleic acid sequences may be introduced in these vectors and the vectors may be introduced in *E. coli* grown in an appropriate medium. Vectors may be recovered from cells upon harvesting and lysing the cells.

Recombinant expression vectors suitable for the introduction of nucleic acid sequences in plant cells include *Agrobacterium* and *Rhizobium* based vectors such as the Ti and Ri plasmids. *Agrobacterium* based vectors typically carry at least one T-DNA border sequence and include vectors such pBIN 19 (Bevan (1984) Nucl Acids Res. Vol. 12, 22:8711-8721) and other binary vector systems (for example: US Patent 4,940,838).

Production of cells comprising a first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins,

-52-

immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and/or a first and/or second thioredoxin-related protein and oil-body-targeting-proteins

In accordance with the present invention, the recombinant expression vectors are introduced into the cell that is selected and the selected cells are grown to produce the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, a first and/or second thioredoxin-related protein; and the oil-body-targeting-protein either directly or in a progeny cell.

Methodologies to introduce recombinant expression vectors into a cell also referred to herein as "transformation" are well known to the art and vary depending on the cell type that is selected. General techniques to transfer the recombinant expression vectors into the cell include electroporation; chemically mediated techniques, for example CaCl_2 mediated nucleic acid uptake; particle bombardment (biolistics); the use of naturally infective nucleic acid sequences for example virally derived nucleic acid sequences or when plant cells are used *Agrobacterium* or *Rhizobium* derived nucleic acid sequences; PEG mediated nucleic acid uptake, microinjection, and the use of silicone carbide whiskers (Kaeppeler et al. (1990) Plant Cell Rep. 9:415-418) all of which may be used herein.

Introduction of the recombinant expression vector into the cell may result in integration of its whole or partial uptake into host cell genome including the chromosomal DNA or the plastid genome. Alternatively the recombinant expression vector may not be integrated into the genome and replicate independently of the host cell's genomic DNA. Genomic integration of the nucleic acid sequence is typically used as it will allow for stable inheritance of the introduced nucleic acid sequences by subsequent generations of cells and the creation of cell, plant or animal lines.

Particular embodiments involve the use of plant cells. Particular plant cells used herein include cells obtainable from Brazil nut (*Bethollletia excelsa*); castor (*Ricinus communis*); coconut (*Cocus nucifera*); coriander (*Coriandrum*

-53-

sativum); cotton (*Gossypium* spp.); groundnut (*Arachis hypogaea*); jojoba (*Simmondsia chinensis*); linseed/flax (*Linum usitatissimum*); maize (*Zea mays*); mustard (*Brassica* spp. and *Sinapis alba*); oil palm (*Elaeis guineensis*); olive (*Olea europaea*); rapeseed (*Brassica* spp.); safflower (*Carthamus tinctorius*); soybean
5 (*Glycine max*); squash (*Cucurbita maxima*); barley (*Hordeum vulgare*); wheat (*Triticum aestivum*) and sunflower (*Helianthus annuus*).

Transformation methodologies for dicotyledonous plant species are well known. Generally *Agrobacterium* mediated transformation is utilized because of its high efficiency as well as the general susceptibility by many, if not all
10 dicotyledonous plant species. *Agrobacterium* transformation generally involves the transfer of a binary vector (e.g. pBIN19) comprising the DNA of interest to an appropriate *Agrobacterium* strain (e.g. CIB542) by for example tri-parental mating with an *E. coli* strain carrying the recombinant binary vector and an *E. coli* strain carrying a helper plasmid capable of mobilization of the binary vector
15 to the target *Agrobacterium* strain, or by DNA transformation of the *Agrobacterium* strain (Hofgen et al. Nucl. Acids. Res. (1988) 16: 9877. Other transformation methodologies that may be used to transform dicotyledonous plant species include biolistics (Sanford (1988) Trends in Biotechn. 6: 299-302); electroporation (Fromm et al. (1985) Proc. Natl. Acad. Sci. USA 82: 5824-
20 5828); PEG mediated DNA uptake (Potrykus et al. (1985) Mol. Gen. Genetics 199: 169-177); microinjection (Reich et al. Bio/Techn. (1986) 4: 1001-1004) and silicone carbide whiskers (Kaepler et al. (1990) Plant Cell Rep. 9: 415-418). The exact transformation methodologies typically vary somewhat depending on the plant species that is used.

25 In a particular embodiment the oil bodies are obtained from safflower and the recombinant proteins are expressed in safflower. Safflower transformation has been described by Baker and Dyer (Plant Cell Rep. (1996) 16: 106-110).

Monocotyledonous plant species may now also be transformed using a variety of methodologies including particle bombardment (Christou et al. (1991) Biotechn. 9: 957-962; Weeks et al. Plant Physiol. (1993) 102: 1077-1084;
30 Gordon-Kamm et al. Plant Cell (1990) 2: 603-618) PEG mediated DNA uptake

-54-

(EP 0 292 435; 0 392 225) or *Agrobacterium*-mediated transformation (Goto-Fumiyuki et al (1999) Nature-Biotech. 17 (3):282-286).

Plastid transformation is described in US Patents 5,451,513; 5,545,817 and 5,545,818; and PCT Patent Applications 95/16783; 98/11235 and
5 00/39313) Basic chloroplast transformation involves the introduction of cloned plastid DNA flanking a selectable marker together with the nucleic acid sequence of interest into a suitable target tissue using for example biolistics or protoplast transformation. Selectable markers that may be used include for example the bacterial *aadA* gene (Svab et al. (1993) Proc. Natl. Acad. Sci. USA 90: 913-
10 917). Plastid promoters that may be used include for example the tobacco *clpP* gene promoter (PCT Patent Application 97/06250).

In another embodiment, the invention chimeric nucleic acid constructs provided herein are directly transformed into the plastid genome. Plastid transformation technology is described extensively in U.S. Patent Nos.
15 5,451,513, 5,545,817, 5,545,818 and 5,576,198; in PCT application nos. WO 95/16783 and WO 97/32977; and in McBride et. al., *Proc Natl Acad Sci USA* 91: 7301-7305 (1994), the entire disclosures of all of which are hereby incorporated by reference. In one embodiment, plastid transformation is achieved via biolistics, first carried out in the unicellular green alga
20 *Chlamydomonas reinhardtii* (Boynton et al. (1988) *Science* 240:1534-1537)) and then extended to *Nicotiana tabacum* (Svab et al. (1990) *Proc Natl Acad Sci USA* 87:8526-8530), combined with selection for cis-acting antibiotic resistance loci (spectinomycin or streptomycin resistance) or complementation of non-photosynthetic mutant phenotypes.

25 In another embodiment, tobacco plastid transformation is carried out by particle bombardment of leaf or callus tissue, or polyethylene glycol (PEG)-mediated uptake of plasmid DNA by protoplasts, using cloned plastid DNA flanking a selectable antibiotic resistance marker. For example, 1 to 1.5 kb flanking regions, termed targeting sequences, facilitate homologous
30 recombination with the plastid genome and allow the replacement or modification of specific regions of the 156 kb tobacco plastid genome. In one embodiment, point mutations in the plastid 16S rDNA and *rps12* genes

-55-

conferring resistance to spectinomycin and/or streptomycin can be utilized as selectable markers for transformation (Svab *et al.* (1990) *Proc Natl Acad Sci USA* 87:8526-8530; Staub *et al.* (1992) *Plant Cell* 4:39-45, the entire disclosures of which are hereby incorporated by reference), resulting in stable

5 homoplasmic transformants at a frequency of approximately one per 100 bombardments of target leaves. The presence of cloning sites between these markers allows creation of a plastid targeting vector for introduction of foreign genes (Staub *et al.* (1993) *EMBO J* 12:601-606, the entire disclosure of which is hereby incorporated by reference). In another embodiment, substantial increases

10 in transformation frequency can be obtained by replacement of the recessive rRNA or r-protein antibiotic resistance genes with a dominant selectable marker, the bacterial *aadA* gene encoding the spectinomycin-detoxifying enzyme aminoglycoside-3'-adenyltransferase (Svab *et al.* (1993) *Proc Natl Acad Sci USA* 90: 913-917, the entire disclosure of which is hereby incorporated by reference).

15 This marker has also been used successfully for high-frequency transformation of the plastid genome of the green alga *Chlamydomonas reinhardtii* (Goldschmidt-Clermont, M. (1991) *Nucl Acids Res* 19, 4083-4089, the entire disclosure of which is hereby incorporated by reference). In other embodiments, plastid transformation of protoplasts from tobacco and the moss *Physcomitrella*

20 can be attained using PEG-mediated DNA uptake (O'Neill *et al.* (1993) *Plant J* 3:729-738; Koop *et al.* (1996) *Planta* 199:193-201, the entire disclosures of which are hereby incorporated by reference).

Both particle bombardment and protoplast transformation are also contemplated for use herein. Plastid transformation of oilseed plants has been

25 successfully carried out in the genera *Arabidopsis* and *Brassica* (Sikdar *et al.* (1998) *Plant Cell Rep* 18:20-24; PCT Application WO 00/39313, the entire disclosures of which are hereby incorporated by reference).

A chimeric nucleic sequence construct is inserted into a plastid expression cassette including a promoter capable of expressing the construct in

30 plant plastids. A particular promoter capable of expression in a plant plastid is, for example, a promoter isolated from the 5' flanking region upstream of the coding region of a plastid gene, which may come from the same or a different

-56-

species, and the native product of which is typically found in a majority of plastid types including those present in non-green tissues. Gene expression in plastids differs from nuclear gene expression and is related to gene expression in prokaryotes (Stern *et al.* (1997) *Trends in Plant Sci* 2:308-315, the entire

5 disclosure of which is hereby incorporated by reference).

Plastid promoters generally contain the -35 and -10 elements typical of prokaryotic promoters, and some plastid promoters called PEP (plastid-encoded RNA polymerase) promoters are recognized by an *E. coli*-like RNA polymerase mostly encoded in the plastid genome, while other plastid promoters called NEP

10 promoters are recognized by a nuclear-encoded RNA polymerase. Both types of plastid promoters are suitable for use herein. Examples of plastid promoters include promoters of clpP genes such as the tobacco clpP gene promoter (WO 97/06250, the entire disclosure of which is hereby incorporated by reference) and the *Arabidopsis* clpP gene promoter (U.S. Application No. 09/038,878, the

15 entire disclosure of which is hereby incorporated by reference). Another promoter capable of driving expression of a chimeric nucleic acid construct in plant plastids comes from the regulatory region of the plastid 16S ribosomal RNA operon (Harris *et al.*, (1994) *Microbiol Rev* 58:700-754; Shinozaki *et al.* (1986) *EMBO J* 5:2043-2049, the entire disclosures of both of which are hereby

20 incorporated by reference). Other examples of promoters capable of driving expression of a nucleic acid construct in plant plastids include a psbA promoter or am rbcL promoter. A plastid expression cassette preferably further includes a plastid gene 3' untranslated sequence (3' UTR) operatively linked to a chimeric nucleic acid construct of the present invention. The role of untranslated

25 sequences is preferably to direct the 3' processing of the transcribed RNA rather than termination of transcription. An exemplary 3' UTR is a plastid rps16 gene 3' untranslated sequence, or the *Arabidopsis* plastid psbA gene 3' untranslated sequence. In a further embodiment, a plastid expression cassette includes a poly-G tract instead of a 3' untranslated sequence. A plastid expression

30 cassette also preferably further includes a 5' untranslated sequence (5' UTR) functional in plant plastids, operatively linked to a chimeric nucleic acid construct provided herein.

-57-

A plastid expression cassette is contained in a plastid transformation vector, which preferably further includes flanking regions for integration into the plastid genome by homologous recombination. The plastid transformation vector may optionally include at least one plastid origin of replication. The present
5 invention also encompasses a plant plastid transformed with such a plastid transformation vector, wherein the chimeric nucleic acid construct is expressible in the plant plastid. Also encompassed herein is a plant or plant cell, including the progeny thereof, including this plant plastid. In a particular embodiment, the plant or plant cell, including the progeny thereof, is homoplasmic for transgenic
10 plastids.

Other promoters capable of driving expression of a chimeric nucleic acid construct in plant plastids include transactivator-regulated promoters, preferably heterologous with respect to the plant or to the subcellular organelle or component of the plant cell in which expression is effected. In these cases, the
15 DNA molecule encoding the transactivator is inserted into an appropriate nuclear expression cassette which is transformed into the plant nuclear DNA. The transactivator is targeted to plastids using a plastid transit peptide. The transactivator and the transactivator-driven DNA molecule are brought together either by crossing a selected plastid-transformed line with and a transgenic line
20 containing a DNA molecule encoding the transactivator supplemented with a plastid-targeting sequence and operably linked to a nuclear promoter, or by directly transforming a plastid transformation vector containing the desired DNA molecule into a transgenic line containing a chimeric nucleic acid construct encoding the transactivator supplemented with a plastid-targeting sequence
25 operably linked to a nuclear promoter. If the nuclear promoter is an inducible promoter, in particular a chemically inducible embodiment, expression of the chimeric nucleic acid construct in the plastids of plants is activated by foliar application of a chemical inducer. Such an inducible transactivator-mediated plastid expression system is preferably tightly regulatable, with no detectable
30 expression prior to induction and exceptionally high expression and accumulation of protein following induction.

-58-

A particular transactivator is, for example, viral RNA polymerase.

Particular promoters of this type are promoters recognized by a single sub-unit RNA polymerase, such as the T7 gene 10 promoter, which is recognized by the bacteriophage T7 DNA-dependent RNA polymerase. The gene encoding the T7
5 polymerase is preferably transformed into the nuclear genome and the T7 polymerase is targeted to the plastids using a plastid transit peptide. Promoters suitable for nuclear expression of a gene, for example a gene encoding a viral RNA polymerase such as the T7 polymerase, are described above and elsewhere in this application. Expression of chimeric nucleic acid constructs in plastids can
10 be constitutive or can be inducible, and such plastid expression can be also organ- or tissue-specific. Examples of various expression systems are extensively described in WO 98/11235, the entire disclosure of which is hereby incorporated by reference. Thus, in one aspect, the present invention utilizes coupled expression in the nuclear genome of a chloroplast-targeted phage T7
15 RNA polymerase under the control of the chemically inducible PR-1a promoter, for example of the PR-1 promoter of tobacco, operably linked with a chloroplast reporter transgene regulated by T7 gene 10 promoter/terminator sequences, for example as described in as in US Patent No. 5,614,395 the entire disclosure of which is hereby incorporated by reference. In another embodiment, when
20 plastid transformants homoplasmic for the maternally inherited TR or NTR genes are pollinated by lines expressing the T7 polymerase in the nucleus, F1 plants are obtained that carry both transgene constructs but do not express them until synthesis of large amounts of enzymatically active protein in the plastids is triggered by foliar application of the PR-1a inducer compound
25 benzo(1,2,3)thiadiazole-7-carbothioic acid S-methyl ester (BTH).

In a particular embodiment, two or more genes, for example TR and NTR genes, are transcribed from the plastid genome from a single promoter in an operon-like polycistronic gene. In one embodiment, the operon-like polycistronic gene includes an intervening DNA sequence between two genes in the operon-
30 like polycistronic gene. In a particular embodiment, the intervening DNA sequence is not present in the plastid genome to avoid homologous recombination with plastid sequences. In another embodiment, the DNA

-59-

sequence is derived from the 5' untranslated (UTR) region of a non-eukaryotic gene, preferably from a viral 5'UTR, preferably from a 5'UTR derived from a bacterial phage, such as a T7, T3 or SP6 phage. In one embodiment, a portion of the DNA sequence may be modified to prevent the formation of RNA
5 secondary structures in an RNA transcript of the operon-like polycistronic gene, for example between the DNA sequence and the RBS of the downstream gene. Such secondary structures may inhibit or repress the expression of the downstream gene, particularly the initiation of translation. Such RNA secondary structures are predicted by determining their melting temperatures using
10 computer models and programs such as the "mfold" program version 3 (available from Zuker and Turner, Washington University School of Medicine, St-Louis, MO) and other methods known to one skilled in the art.

The presence of the intervening DNA sequence in the operon-like polycistronic gene increases the accessibility of the RBS of the downstream
15 gene, thus resulting in higher rates of expression. Such strategy is applicable to any two or more genes to be transcribed from the plastid genome from a single promoter in an operon-like chimeric heteromultimeric gene.

Following transformation the cells are grown, typically in a selective medium allowing the identification of transformants. Cells may be harvested in
20 accordance with methodologies known to the art. In order to associate the oil bodies with the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, and a first and/or second thioredoxin-related
25 protein, the integrity of cells may be disrupted using any physical, chemical or biological methodology capable of disrupting the cells' integrity. These methodologies are generally cell-type dependent and known to the skilled artisan. Where plants are employed they may be regenerated into mature plants using plant tissue culture techniques generally known to the skilled artisan.
30 Seeds may be harvested from mature transformed plants and used to propagate the plant line. Plants may also be crossed and in this manner, contemplated herein is the breeding of cell lines and transgenic plants that vary in genetic

-60-

background. It is also possible to cross a plant line comprising the first recombinant polypeptide with a plant line comprising the second recombinant polypeptide. Accordingly, also provided herein are methods of producing in a plant a recombinant multimeric-protein-complex, said method comprising:

- 5 (a) preparing a first plant comprising cells, said cells comprising oil bodies and a first recombinant polypeptide, such as a redox protein (e.g., a thioredoxin-related protein, and the like) or an immunoglobulin-polypeptide-chain, wherein said first recombinant polypeptide is capable of associating with said oil bodies through an oil-body-targeting-protein;
- 10 (b) preparing a second plant comprising cells, said cells comprising oil bodies and a second recombinant polypeptide, such as a second redox protein (e.g., a thioredoxin-related protein, and the like) or a second immunoglobulin-polypeptide-chain; and
- (c) sexually crossing said first plant with said second plant to produce a progeny
- 15 plant comprising cells, said cells comprising oil bodies, wherein said oil bodies are capable of associating with said first recombinant polypeptide, and said first recombinant recombinant polypeptide is capable of associating with said second recombinant polypeptide to form said recombinant multimeric-protein-complex.

- The second recombinant polypeptide may also associate with the oil
- 20 bodies. Accordingly, also provided herein are methods of producing in a plant a recombinant multimeric-protein-complex, said method comprising:
 - (a) preparing a first plant comprising cells, said cells comprising oil bodies and a first recombinant polypeptide, such as a redox (or thioredoxin-related) protein or immunoglobulin-polypeptide-chain, wherein said first recombinant polypeptide is
 - 25 capable of associating with said oil bodies through an oil-body-targeting-protein;
 - (b) preparing a second plant comprising cells, said cells comprising oil bodies and a second recombinant polypeptide, such as a second redox (thioredoxin-related) protein or a second immunoglobulin-polypeptide-chain, wherein said second recombinant polypeptide is capable of associating with said oil bodies through an
 - 30 oil body targeting protein; and
 - (c) sexually crossing said first plant with said second plant to produce a progeny plant comprising cells, said cells comprising oil bodies, wherein said oil bodies

-61-

are capable of associating with said first recombinant polypeptide, and said first recombinant recombinant polypeptide is capable of associating with said second recombinant polypeptide to form said recombinant multimeric-protein-complex.

Isolation of Oil bodies

5 The oil bodies provided herein may be obtained from any cell containing oil bodies, including any animal cell; plant cell; fungal cell; for example a yeast cell, algae cell; or bacterial cell. Any process suitable for the isolation oil bodies from cells may be used herein. Processes for the isolation of oil bodies from plant seed cells have been described in US Patents (6,146,645 and 6,183,762)
10 and the isolation of oil bodies from yeast cells has been described by Ting et al. (1997) J. Biol. Chem. 272: 3699-3706).

 In certain embodiments, the oil bodies are obtained from a plant cell such as for example a pollen cell; a fruit cell; a spore cell; a nut cell; mesocarp cell; for example the mesocarp cells obtainable from olive (*Olea europaea*) or avocado
15 (*Persea americana*); or a seed cell. In particular embodiments the oil bodies are obtained from a plant seed cell. The seeds can be obtained from a transgenic plant according to the present invention. In particular embodiments, a seed of a transgenic plant according to the present invention contains the first and/or
20 heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or first and/or second thioredoxin-related proteins in a concentration of at least about 0.5% of total cellular seed protein. In further embodiments, a seed of a transgenic plant provided herein contains a
25 recombinant polypeptide or multimeric-protein-complex in a concentration of at least about 0.5%, 0.6%, 0.7%, 0.8%, 0.9%, 1.0%, 1.25%, 1.5%, 1.75%, 2.0%, 2.25%, 2.5%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10% or more, of total cellular seed protein. The upper limits of the recombinant polypeptide or multimeric-protein-complex concentration can be up to about 8%, 9%, 10%,
30 11%, 12%, 13%, 14%, 15%. Thus, the ranges at least about 0.5% up to about 15%; at least about 1.0% up to about 10%; and at least about 5% up to about 8% are among the various ranges contemplated herein.

-62-

Among the plant seeds useful in this regard are plant seeds obtainable from the group of plant species consisting of Brazil nut (*Betholletia excelsa*); castor (*Ricinus communis*); coconut (*Cocus nucifera*); coriander (*Coriandrum sativum*); cotton (*Gossypium* spp.); groundnut (*Arachis hypogaea*); jojoba
5 (*Simmondsia chinensis*); linseed/flax (*Linum usitatissimum*); maize (*Zea mays*); mustard (*Brassica* spp. and *Sinapis alba*); oil palm (*Elaeis guineensis*); olive (*Olea europaea*); rapeseed (*Brassica* spp.); safflower (*Carthamus tinctorius*); soybean (*Glycine max*); squash (*Cucurbita maxima*); sunflower (*Helianthus annuus*); barley (*Hordeum vulgare*); wheat (*Triticum aestivum*) and mixtures thereof. In
10 a particular embodiment, oil bodies are obtainable from the seeds obtainable from safflower (*Carthamus tinctorius*).

In order to prepare oil bodies from plant seeds, plants are grown and allowed to set seed in accordance with common agricultural practices. Thus, the present invention also provides seeds comprising oil bodies, wherein said oil
15 bodies further comprise invention multimeric-protein-complexes described herein. Upon harvesting the seed and, if necessary the removal of large insoluble materials such as stones or seed hulls, by for example sieving or rinsing, any process suitable for the isolation of oil bodies from seeds may be used herein. A typical process involves grinding of the seeds followed by an aqueous extraction
20 process.

Seed grinding may be accomplished by any comminuting process resulting in a substantial disruption of the seed cell membrane and cell walls without compromising the structural integrity of the oil bodies present in the seed cell. Suitable grinding processes in this regard include mechanical pressing
25 and milling of the seed. Wet milling processes such as described for cotton (Lawhon et al. (1977) J. Am. Oil Chem. Soc. 63: 533-534) and soybean (US Patent 3,971,856; Carter et al. (1974) J. Am. Oil Chem. Soc. 51: 137-141) are particularly useful in this regard. Suitable milling equipment capable of industrial scale seed milling include colloid mills, disc mills, pin mills, orbital mills, IKA mills
30 and industrial scale homogenizers. The selection of the milling equipment will depend on the seed, which is selected, as well as the throughput requirement.

Solid contaminants such as seed hulls, fibrous materials, undissolved carbohydrates, proteins and other insoluble contaminants are subsequently preferably removed from the ground seed fraction using size exclusion based methodologies such as filtering or gravitational based methods such as a centrifugation based separation process. Centrifugation may be accomplished using for example a decantation centrifuge such as a HASCO 200 2-phase decantation centrifuge or an NX310B (Alpha Laval). Operating conditions are selected such that a substantial portion of the insoluble contaminants and sediments and may be separated from the soluble fraction.

Following the removal of insolubles the oil body fraction may be separated from the aqueous fraction. Gravitational based methods as well as size exclusion based technologies may be used. Gravitational based methods that may be used include centrifugation using for example a tubular bowl centrifuge such as a Sharples AS-16 or AS-46 (Alpha Laval), a disc stack centrifuge or a hydrocyclone, or separation of the phases under natural gravitation. Size exclusion methodologies that may be used include membrane ultra filtration and crossflow microfiltration.

Separation of solids and separation of the oil body phase from the aqueous phase may also be carried out concomitantly using gravity based separation methods or size exclusion based methods.

The oil body preparations obtained at this stage in the process are generally relatively crude and depending on the application of the oil bodies, it may be desirable to remove additional contaminants. Any process capable of removing additional seed contaminants may be used in this regard. Conveniently the removal of these contaminants from the oil body preparation may be accomplished by resuspending the oil body preparation in an aqueous phase and re-centrifuging the resuspended fraction, a process referred to herein as "washing the oil bodies". The washing conditions selected may vary depending on the desired purity of the oil body fractions. For example where oil bodies are used in pharmaceutical compositions, generally a higher degree of purity may be desirable than when the oil bodies are used in food preparations. The oil bodies may be washed one or more times depending on the desired purity and the ionic

-64-

strength, pH and temperature may all be varied. Analytical techniques may be used to monitor the removal of contaminants. For example SDS gel electrophoresis may be employed to monitor the removal of seed proteins.

The entire oil body isolation process may be performed in a batch wise fashion or continuous flow. In a particular embodiment, industrial scale continuous flow processes are utilized.

Through the application of these and similar techniques the skilled artisan is able to obtain oil bodies from any cell comprising oil bodies. The skilled artisan will recognize that generally the process will vary somewhat depending on the cell type that is selected. However, such variations may be made without departing from the scope and spirit of the present invention.

Association of the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, the first and/or second thioredoxin-related proteins with oil bodies.

In accordance with the present invention, the oil bodies are associated with either the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, the first and/or second thioredoxin-related proteins through association with an oil-body-targeting-protein capable of association with these multimeric-protein-complexes and the oil bodies. As used herein the phrase "associating the oil bodies with the multimeric-protein-complex" means that the oil bodies are brought in proximity of the multimeric-protein-complexes in a manner that allows the association of the oil bodies with either the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins. The association of the oil bodies with the multimeric-protein-complexes is accomplished by association of the oil-body-targeting-protein with

both the oil body and with the multimeric-protein-complex. In particular embodiments, the cells expressing the multimeric-protein-complex associate with the oil bodies that are obtainable from these same cells, which permits the convenient production and isolation of the multimeric-protein-complex, including

5 the first and/or second recombinant polypeptides, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins, in an oil body-comprising host cell system. Accordingly, in one embodiment, the association of the oil

10 body with the multimeric-protein-complex is accomplished intracellularly during the growth of the cell. For example, a redox fusion polypeptide may be fused to an oil-body-protein and the chimeric protein may be expressed in oil body-containing plant seeds. Isolation of the oil bodies from the seeds in this case results in isolation of oil bodies comprising either the first and/or second

15 recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins. In another embodiment, in which the multimeric-protein-complex associates with oil bodies obtainable

20 from the same cells in which the complex is produced, the association of the oil bodies with the multimeric-protein-complex is accomplished upon disrupting the cell's integrity.

For example, the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-

25 fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins may be expressed in such a manner that it is targeted to the endomembrane system of the seed cells. Oil bodies present in the same seed cells comprising an oil-body-targeting-protein capable of

30 association with these multimeric-protein-complexes, for example an oleosin linked to a single chain antibody capable of association with a recombinant polypeptide or multimeric-protein-complex, may then associate with the

-66-

recombinant polypeptide or multimeric-protein-complex upon grinding of the seed.

In accordance with this embodiment, plant seed cells comprising a light and heavy chain of an immunoglobulin targeted to the plant apoplast can be prepared. These particular seed cells are prepared to further comprise oil bodies associated with an oil-body-targeting-protein capable of association with the immunoglobulin, such as for example, an oleosin-protein A fusion protein, and the like. Upon grinding of the seed, the oil bodies comprising protein A associate with the immunoglobulin through binding.

10 In yet another embodiment, the oil bodies used to associate with the multimeric-protein-complex are obtained from a cellular source different from the cell comprising the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-chains, redox-fusion-polypeptides, or the first and/or second
15 thioredoxin-related proteins, such as from a separate plant line. For example, oil bodies associated with protein A may be prepared from one plant line. These oil bodies may then be mixed with ground seeds comprising an apoplastically expressed light and heavy chain constituting an immunoglobulin. Alternatively, a
20 plant line comprising oil bodies associated with protein A may be crossed with a plant line comprising an immunoglobulin.

The first recombinant polypeptide, second recombinant polypeptide and oil-body-targeting-protein may also be prepared in separate cellular compartments. Association of the first polypeptide, second polypeptide, and oil
25 body then may occur upon disruption of the cell's integrity. For example, various mechanisms for targeting gene products are known to exist in plants, and the sequences controlling the functioning of these mechanisms have been characterized in some detail. For example, the targeting of gene products to the chloroplast is controlled by a transit sequence found at the amino terminal end of
30 various proteins which is cleaved during chloroplast import to yield the mature protein (Comai *et al.* (1988) *J Biol Chem* 263: 15104-15109). Other gene products are localized to other organelles such as the mitochondrion and the

-67-

peroxisome (Unger *et al.* (1989) *Plant Mol Biol* 13:411-418). The cDNAs encoding these products can be manipulated to target heterologous gene products to these organelles. In addition, sequences have been characterized which cause the targeting of gene products to other cell compartments.

- 5 Amino terminal sequences are responsible for targeting to the ER, the apoplast, and extracellular secretion from aleurone cells (Koehler & Ho (1990) *Plant Cell* 2:769-783). Additionally, amino terminal sequences in conjunction with carboxy terminal sequences are responsible for vacuolar targeting of gene products (Shinshi *et al.*, (1990) *Plant Mol Biol* 14:357-368). By the fusion of the
- 10 appropriate targeting sequences described above to transgene sequences of interest it is possible to direct the transgene product to the desired organelle or cell compartment.

As hereinbefore mentioned, the redox protein obtained using the methods provided herein is enzymatically active while associated with the oil body.

- 15 Preferably the redox protein is at least 5 times more active when produced as a redox fusion polypeptide with a second redox protein relative to its production in association with an oil body as a non-fusion polypeptide (i.e. without the second redox protein). More preferably the redox protein is at least 10 times more active when produced as a redox fusion polypeptide.

- 20 The activity of the redox fusion polypeptide may be determined in accordance with methodologies generally known to the art (see for example: Johnson et al (1984) *J. of Bact.* Vol. 158 3:1061-1069) and may be optimized by for example the addition of detergents, including ionic and non-ionic detergents.

25 **Formulation of Oil Bodies**

- In accordance with a particular embodiment, the oil bodies comprising the first and/or second recombinant polypeptides, multimeric-protein-complexes, heteromultimeric-protein-complexes, multimeric-fusion-proteins, heteromultimeric-fusion-proteins, immunoglobulins, immunoglobulin-polypeptide-
- 30 chains, redox-fusion-polypeptides, or the first and/or second thioredoxin-related proteins, are preferably formulated into an emulsion. The emulsion is preferably used in the preparation of a pharmaceutical composition, personal care or a food

-68-

product. In emulsified form, the oil body offers certain desirable properties, such as for example excellent compatibility with the human skin.

5 In particular embodiments, the oil body formulation is stabilized so that a final product may be obtained which may be stored and preserved for longer periods of time. As used herein, the term "stabilized oil body preparation" refers to an oil body preparation that is prepared so that the formulation does not undergo undesirable physical or chemical alterations when the oil body preparation is stored. The stabilization requirements may vary depending on the final product. For example personal care products are preferably stable for at
10 least one year at room temperature while additionally being able to withstand short temperature fluctuations. Pharmaceutical formulations may in some cases be less stable as they may be stored at lower temperatures thereby preventing the occurrence of undesirable reactions.

 In general, stabilization techniques that may be used herein include any
15 and all methods for the preservation of biological material including the addition of chemical agents, temperature modulation based methodologies, radiation-based technologies and combinations thereof. In particular embodiments small amounts of stabilizing chemical agents are mixed with the oil body formulation to achieve stabilization. These chemical agents include *inter alia* preservatives,
20 antioxidants, acids, salts, bases, viscosity modifying agents, emulsifiers, gelling agents and mixtures thereof and may all be used to stabilize the oil body preparation. In view of the presence of the redox fusion polypeptide the stabilizing agent is generally selected to be compatible with and resulting in good enzymatic function of the redox fusion polypeptide.

25 Diagnostic parameters to assess the stability of the oil body preparation may be as desired and include all parameters indicative of undesirable qualitative or quantitative changes with respect to chemical or physical stability. Typical parameters to assess the oil body preparation over time include color, odor, viscosity, texture, pH and microbial growth, and enzymatic activity.

30 In particular embodiments, the oil body formulation is stabilized prior to the addition of further ingredients that may be used to prepare the final product. However, in other embodiments, it is nevertheless possible to formulate the

final formulation using non-stabilized oil bodies and stabilize the final formulation.

The final preparations may be obtained using one or more additional ingredients and any formulation process suitable for the preparation of a formulation comprising oil bodies. Ingredients and processes employed will
5 generally vary depending on the desired use of the final product, will be art recognized and may be as desired. Ingredients and processes that may be used herein include those described in US Patents (US Patents 6,146,645 and 6,183,762) which are incorporated by reference herein.

In particular embodiments, the redox fusion polypeptide comprises a
10 thioredoxin and a thioredoxin-reductase. Accordingly, provided herein are oil bodies comprising a thioredoxin/thioredoxin-reductase fusion polypeptide. Also provided herein is a formulation containing oil bodies comprising a thioredoxin/thioredoxin-reductase fusion capable of treating or protecting a target against oxidative stress. The stress of the target is treated or prevented by
15 contacting the target with the formulation. The target may be any substance susceptible to oxidative stress, including any molecule, molecular complex, cell, tissue or organ.

In another embodiment, provided herein is a formulation containing oil bodies comprising a thioredoxin/thioredoxin-reductase fusion capable of
20 chemically reducing a target. Contacting the target with the formulation reduces the target. The target may be any substance susceptible to reduction, including any molecule or molecular complex. Particularly susceptible targets in this regard are the disulfide bonds present in proteins.

The oil bodies comprising thioredoxin/thioredoxin-reductase may be used
25 to prepare formulations used to reduce the allergenicity of food or increase the digestibility of food. Preferably, the method of reducing the food allergenicity is practiced by mixing the thioredoxin/thioredoxin-reductase comprising oil bodies with food or food ingredients selected from a variety of sources including for example wheat flour, wheat dough, milk, cheese, soya, yogurt and ice cream.
30 The thioredoxin/thioredoxin-reductase comprising oil bodies may also be used to increase the digestibility of milk as well as other disulfide containing proteins (Jiao, J. et al. (1992) J. Agric. Food Chem 40: 2333-2336). Further food

-70-

applications include the use of the oil thioredoxin/thioredoxin-reductase comprising oil bodies as a food additive to enhance dough strength and bread quality properties (Wong et al., (1993) J. Cereal Chem. 70: 113-114; Kobrehel et al. (1994) Gluten Proteins: Association of Cereal Research; Detmold, Germany).

Also provided herein are pharmaceutical compositions comprising, in a pharmaceutically active carrier: oil bodies comprising a thioredoxin/thioredoxin-reductase; oil bodies comprising multimeric-protein-complexes, such as heteromultimeric-protein-complexes; isolated thioredoxin/thioredoxin-reductase fusion proteins; or isolated multimeric-protein-complexes. These pharmaceutical compositions may be used for the treatment of reperfusion injury (Aota et al. (1996) J. Cardio. Pharmacol. (1996) 27: 727-732), cataracts (US Patent US 4,771,036), chronic obstructive pulmonary disease (COPD) (MacNee et al. (1999) Am. J. Respir. Crit. Care Med. 160:S58-S65), diabetes (Hotta et al. J. Exp. Med. 188: 1445-1451), envenomation (PCT Patent Application 99/20122; US Patent 5,792,506), bronchiopulmonary disease (MacNee (2000) Chest 117:3035-3175); malignancies (PCT Patent Application 91/04320) and the alleviation of the allergenic potential of airborne, for example pollen- derived, and contact allergens (PCT Patent Application 00/44781). Other diseases or conditions that may be treated with the pharmaceutical compositions provided herein include: psoriasis, wound healing, sepsis, GI bleeding, intestinal bowel disease (IBD), ulcers, transplantation, GERD (gastro esophageal reflux disease).

In another embodiment, the pharmaceutical compositions provided herein, particularly those comprising one or more redox proteins alone or in combination with oil bodies, can be used in the treatment of inflammatory and viral diseases by reductively inactivating phospholipase A2, one of the contributing factors in inflammatory diseases. Additionally, the redox fusion polypeptide system has been found to function as a self-defense mechanism in response to environmental stimuli, including oxidative stress caused by UV-generated free radicals. Consequently, redox proteins, e.g., oleosin-thioredoxin, oleosin-thioredoxin-reductase, the various redox fusion polypeptides described herein, provide beneficial effects in certain skin conditions such as psoriasis, skin

-71-

cancer, dandruff, diaper rash, dermatitis, acne, sun damage, aging, inflammation, and the like.

In another embodiment, oil-body-thioredoxin-related fusion proteins, e.g., oleosin-Thioredoxin-reductase, can also be used as a venom antidote. Many
5 animal venoms and other toxins contain disulfide bonds, including all snake venom neurotoxins, some bacterial neurotoxins including tetanus and botulinum A, bee venom phospholipase A₂, and scorpion venom. In a further embodiment, the redox protein related pharmaceutical compositions provided herein can be used to inactivate venom toxins by reduction of disulfide bonds. A method of
10 treating an individual suffering from the effects of a venom or toxin can include the step of administering an effective dose of a pharmaceutical composition, in a pharmaceutically effective carrier in an amount sufficient to relieve or reverse the effects of the venom toxin on the individual.

The pharmaceutical compositions provided herein are preferably
15 formulated for single dosage administration. The concentrations of the compounds in the formulations are effective for delivery of an amount, upon administration, that is effective for the intended treatment. Typically, the compositions are formulated for single dosage administration. To formulate a composition, the weight fraction of a compound or mixture thereof is dissolved,
20 suspended, dispersed or otherwise mixed in a selected vehicle at an effective concentration such that the treated condition is relieved or ameliorated. Pharmaceutical carriers or vehicles suitable for administration of the compounds provided herein include any such carriers known to those skilled in the art to be suitable for the particular mode of administration.

25 In addition, the compounds may be formulated as the sole pharmaceutically active ingredient in the composition or may be combined with other active ingredients. Liposomal suspensions, including tissue-targeted liposomes, may also be suitable as pharmaceutically acceptable carriers. These may be prepared according to methods known to those skilled in the art. For
30 example, liposome formulations may be prepared as described in U.S. Patent No. 4,522,811.

-72-

The active compound is included in the pharmaceutically acceptable carrier in an amount sufficient to exert a therapeutically useful effect in the absence of undesirable side effects on the patient treated. The therapeutically effective concentration may be determined empirically by testing the compounds
5 in known in vitro and in vivo systems, such as the assays provided herein.

The concentration of active compound in the drug composition will depend on absorption, inactivation and excretion rates of the active compound, the physicochemical characteristics of the compound, the dosage schedule, and amount administered as well as other factors known to those of skill in the art.
10 Typically a therapeutically effective dosage is contemplated. The amounts administered may be on the order of 0.001 to 1 mg/ml, preferably about 0.005-0.05 mg/ml, more preferably about 0.01 mg/ml, of blood volume. Pharmaceutical dosage unit forms are prepared to provide from about 1 mg to about 1000 mg and preferably from about 10 to about 500 mg, more preferably
15 about 25-75 mg of the essential active ingredient or a combination of essential ingredients per dosage unit form. The precise dosage can be empirically determined.

The active ingredient may be administered at once, or may be divided into a number of smaller doses to be administered at intervals of time. It is
20 understood that the precise dosage and duration of treatment is a function of the disease being treated and may be determined empirically using known testing protocols or by extrapolation from in vivo or in vitro test data. It is to be noted that concentrations and dosage values may also vary with the severity of the condition to be alleviated. It is to be further understood that for any particular
25 subject, specific dosage regimens should be adjusted over time according to the individual need and the professional judgment of the person administering or supervising the administration of the compositions, and that the concentration ranges set forth herein are exemplary only and are not intended to limit the scope or use of the claimed compositions and combinations containing them.

30 Preferred pharmaceutically acceptable derivatives include acids, salts, esters, hydrates, solvates and prodrug forms. The derivative is typically selected

-73-

such that its pharmacokinetic properties are superior to the corresponding neutral compound.

Thus, effective concentrations or amounts of one or more of the compounds provided herein or pharmaceutically acceptable derivatives thereof
5 are mixed with a suitable pharmaceutical carrier or vehicle for systemic, topical or local administration to form pharmaceutical compositions. Compounds are included in an amount effective for ameliorating or treating the disorder for which treatment is contemplated. The concentration of active compound in the composition will depend on absorption, inactivation, excretion rates of the active
10 compound, the dosage schedule, amount administered, particular formulation as well as other factors known to those of skill in the art.

Solutions or suspensions used for parenteral, intradermal, subcutaneous, or topical application can include any of the following components: a sterile diluent, such as water for injection, saline solution, fixed oil, polyethylene glycol,
15 glycerine, propylene glycol or other synthetic solvent; antimicrobial agents, such as benzyl alcohol and methyl parabens; antioxidants, such as ascorbic acid and sodium bisulfite; chelating agents, such as ethylenediaminetetraacetic acid (EDTA); buffers, such as acetates, citrates and phosphates; and agents for the adjustment of tonicity such as sodium chloride or dextrose. Parenteral
20 preparations can be enclosed in ampules, disposable syringes or single or multiple dose vials made of glass, plastic or other suitable material.

In instances in which the compounds exhibit insufficient solubility, methods for solubilizing compounds may be used. Such methods are known to those of skill in this art, and include, but are not limited to, using cosolvents,
25 such as dimethylsulfoxide (DMSO), using surfactants, such as Tween®, or dissolution in aqueous sodium bicarbonate. Derivatives of the compounds, such as prodrugs of the compounds may also be used in formulating effective pharmaceutical compositions. For ophthalmic indications, the compositions are formulated in an ophthalmically acceptable carrier. For the ophthalmic uses
30 herein, local administration, either by topical administration or by injection is preferred. Time release formulations are also desirable. Typically, the

-74-

compositions are formulated for single dosage administration, so that a single dose administers an effective amount.

Upon mixing or addition of the compound with the vehicle, the resulting mixture may be a solution, suspension, emulsion or other composition. The form
5 of the resulting mixture depends upon a number of factors, including the intended mode of administration and the solubility of the compound in the selected carrier or vehicle. If necessary, pharmaceutically acceptable salts or other derivatives of the compounds are prepared.

The compound is included in the pharmaceutically acceptable carrier in an
10 amount sufficient to exert a therapeutically useful effect in the absence of undesirable side effects on the patient treated. It is understood that number and degree of side effects depends upon the condition for which the compounds are administered. For example, certain toxic and undesirable side effects are tolerated when treating life-threatening illnesses that would not be tolerated
15 when treating disorders of lesser consequence.

The compounds can also be mixed with other active materials, that do not impair the desired action, or with materials that supplement the desired action known to those of skill in the art. The formulations of the compounds and agents for use herein include those suitable for oral, rectal, topical,
20 inhalational, buccal (*e.g.*, sublingual), parenteral (*e.g.*, subcutaneous, intramuscular, intradermal, or intravenous), transdermal administration or any route. The most suitable route in any given case will depend on the nature and severity of the condition being treated and on the nature of the particular active compound which is being used. The formulations are provided for administration
25 to humans and animals in unit dosage forms, such as tablets, capsules, pills, powders, granules, sterile parenteral solutions or suspensions, and oral solutions or suspensions, and oil-water emulsions containing suitable quantities of the compounds or pharmaceutically acceptable derivatives thereof. The pharmaceutically therapeutically active compounds and derivatives thereof are
30 typically formulated and administered in unit-dosage forms or multiple-dosage forms. Unit-dose forms as used herein refers to physically discrete units suitable for human and animal subjects and packaged individually as is known in the art.

-75-

Each unit-dose contains a predetermined quantity of the therapeutically active compound sufficient to produce the desired therapeutic effect, in association with the required pharmaceutically acceptable carrier, vehicle or diluent.

Examples of unit-dose forms include ampoules and syringes and individually
5 packaged tablets or capsules. Unit-dose forms may be administered in fractions or multiples thereof. A multiple-dose form is a plurality of identical unit-dosage forms packaged in a single container to be administered in segregated unit-dose form. Examples of multiple-dose forms include vials, bottles of tablets or capsules or bottles of pints or gallons. Hence, multiple dose form is a multiple of
10 unit-doses which are not segregated in packaging.

The composition can contain along with the active ingredient: a diluent such as lactose, sucrose, dicalcium phosphate, or carboxymethylcellulose; a lubricant, such as magnesium stearate, calcium stearate and talc; and a binder such as starch, natural gums, such as gum acaciagelatin, glucose, molasses,
15 polivinylypyrrolidine, celluloses and derivatives thereof, povidone, crospovidones and other such binders known to those of skill in the art. Liquid pharmaceutically administrable compositions can, for example, be prepared by dissolving, dispersing, or otherwise mixing an active compound as defined above and optional pharmaceutical adjuvants in a carrier, such as, for example, water,
20 saline, aqueous dextrose, glycerol, glycols, ethanol, and the like, to thereby form a solution or suspension. If desired, the pharmaceutical composition to be administered may also contain minor amounts of nontoxic auxiliary substances such as wetting agents, emulsifying agents, or solubilizing agents, pH buffering agents and the like, for example, acetate, sodium citrate, cyclodextrine
25 derivatives, sorbitan monolaurate, triethanolamine sodium acetate, triethanolamine oleate, and other such agents. Methods of preparing such dosage forms are known, or will be apparent, to those skilled in this art (see, *e.g.*, Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton, Pa., 15th Edition, 1975). The composition or formulation to be administered will
30 contain a quantity of the active compound in an amount sufficient to alleviate the symptoms of the treated subject.

-76-

Dosage forms or compositions containing active ingredient in the range of 0.005% to 100% with the balance made up from non-toxic carrier may be prepared. For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with

5 pharmaceutically acceptable excipients such as binding agents (*e.g.*, pregelatinized maize starch, polyvinyl pyrrolidone or hydroxypropyl methylcellulose); fillers (*e.g.*, lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (*e.g.*, magnesium stearate, talc or silica); disintegrants (*e.g.*, potato starch or sodium starch glycolate); or wetting agents

10 (*e.g.*, sodium lauryl sulphate). The tablets may be coated by methods well-known in the art.

The pharmaceutical preparation may also be in liquid form, for example, solutions, syrups or suspensions, or may be presented as a drug product for reconstitution with water or other suitable vehicle before use. Such liquid

15 preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (*e.g.*, sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (*e.g.*, lecithin or acacia); non-aqueous vehicles (*e.g.*, almond oil, oily esters, or fractionated vegetable oils); and preservatives (*e.g.*, methyl or propyl-p-hydroxybenzoates or

20 sorbic acid).

Formulations suitable for rectal administration are preferably presented as unit dose suppositories. These may be prepared by admixing the active compound with one or more conventional solid carriers, for example, cocoa butter, and then shaping the resulting mixture.

25 Formulations suitable for topical application to the skin or to the eye preferably take the form of an ointment, cream, lotion, paste, gel, spray, aerosol and oil. Carriers which may be used include vaseline, lanoline, polyethylene glycols, alcohols, and combinations of two or more thereof. The topical formulations may further advantageously contain 0.05 to 15 percent by weight

30 of thickeners selected from among hydroxypropyl methyl cellulose, methyl cellulose, polyvinylpyrrolidone, polyvinyl alcohol, poly (alkylene glycols), poly(hydroxyalkyl, (meth)acrylates or poly(meth)acrylamides. A topical

-77-

formulation is often applied by instillation or as an ointment into the conjunctival sac. It can also be used for irrigation or lubrication of the eye, facial sinuses, and external auditory meatus. It may also be injected into the anterior eye chamber and other places. The topical formulations in the liquid state may be
5 also present in a hydrophilic three-dimensional polymer matrix in the form of a strip, contact lens, and the like from which the active components are released.

For administration by inhalation, the compounds for use herein can be delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane,
10 trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol, the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin, for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as
15 lactose or starch.

Formulations suitable for buccal (sublingual) administration include, for example, lozenges containing the active compound in a flavored base, usually sucrose and acacia or tragacanth; and pastilles containing the compound in an inert base such as gelatin and glycerin or sucrose and acacia.

20 The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with an added preservative. The compositions may be suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory
25 agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for reconstitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water or other solvents, before use.

Formulations suitable for transdermal administration may be presented as discrete patches adapted to remain in intimate contact with the epidermis of the
30 recipient for a prolonged period of time. Such patches suitably contain the active compound as an optionally buffered aqueous solution of, for example, 0.1 to 0.2 M concentration with respect to the active compound. Formulations

-78-

suitable for transdermal administration may also be delivered by iontophoresis (see, e.g., *Pharmaceutical Research* 3 (6), 318 (1986)) and typically take the form of an optionally buffered aqueous solution of the active compound.

The pharmaceutical compositions may also be administered by controlled
5 release means and/or delivery devices (see, e.g., in U.S. Patent Nos. 3,536,809; 3,598,123; 3,630,200; 3,845,770; 3,847,770; 3,916,899; 4,008,719; 4,687,610; 4,769,027; 5,059,595; 5,073,543; 5,120,548; 5,354,566; 5,591,767; 5,639,476; 5,674,533 and 5,733,566).

Desirable blood levels may be maintained by a continuous infusion of the
10 active agent as ascertained by plasma levels. It should be noted that the attending physician would know how to and when to terminate, interrupt or adjust therapy to lower dosage due to toxicity, or bone marrow, liver or kidney dysfunctions. Conversely, the attending physician would also know how to and when to adjust treatment to higher levels if the clinical response is not adequate
15 (precluding toxic side effects).

The efficacy and/or toxicity of the pharmaceutical compositions provided herein, alone or in combination with other agents can also be assessed by the methods known in the art (See generally, O'Reilly, *Investigational New Drugs*, 15:5-13 (1997)).

20 The active compounds or pharmaceutically acceptable derivatives may be prepared with carriers that protect the compound against rapid elimination from the body, such as time release formulations or coatings.

Kits containing the compositions and/or the combinations with instructions for administration thereof are provided. The kit may further include
25 a needle or syringe, preferably packaged in sterile form, for injecting the complex, and/or a packaged alcohol pad. Instructions are optionally included for administration of the active agent by a clinician or by the patient.

Finally, the pharmaceutical compositions provided herein containing any of the preceding agents may be packaged as articles of manufacture containing
30 packaging material, a compound or suitable derivative thereof provided herein, which is effective for treatment of a diseases or disorders contemplated herein, within the packaging material, and a label that indicates that the compound or a

-79-

suitable derivative thereof is for treating the diseases or disorders contemplated herein. The label can optionally include the disorders for which the therapy is warranted.

- Also provided herein are personal care formulations containing oil bodies
- 5 comprising a thioredoxin/thioredoxin-reductase fusion polypeptide. Personal care products comprising thioredoxin and thioredoxin-reductase are disclosed in for example Japanese Patent Applications JP9012471A2, JP103743A2, and JP1129785A2 Personal care formulations that may be prepared in accordance with the present invention include formulations capable of improving the physical
- 10 appearance of skin exposed to detrimental environmental stimuli resulting in oxidative stress for example oxidative stress caused by UV-generated free-radicals. The oil bodies comprising thioredoxin/thioredoxin-reductase may also be used to prepare hair care products as described in US Patent Nos. 4,935,231 and 4,973,475 (incorporated herein by reference in their entirety).
- 15 The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

EXAMPLE 1

Isolation of thioredoxin and NADPH thioredoxin-reductase genes

- An *Arabidopsis* silique cDNA library CD4-12 was obtained from the
- 20 *Arabidopsis* Biological Resource Centre (ABRC, <http://aims.cps.msu.edu>) *Arabidopsis* stock centre and used as a template for the isolation of the thioredoxin h (Trxh) and thioredoxin-reductase genes from *Arabidopsis*. For the isolation of the Trxh gene the following primers were synthesized:

GVR833: 5' TACCATGGCTTCGGAAGAAGGA 3' (SEQ ID NO:1)

- 25 The sequence identical to the 5' end of the Trxh gene as published in Rivera-Madrid et al, (1993) Plant Physiol 102: 327-328, is indicated in bold. Underlined is an NcoI restriction site to facilitate cloning. GVR834: 5' GAAAGCTTAAGCCAAGTGTGTTG 3' (SEQ ID NO:2)

- 30 The sequence complementary to the 3' end of the Trxh gene as published in Rivera-Madrid et al, (1993) Plant Physiol 102: 327-328, is indicated in bold. Underlined is an HindIII restriction site to facilitate cloning.

-80-

A Polymerase Chain Reaction (PCR) was carried out using GVR833 and GVR834 as primers and the cDNA library CD4-12 as a template. The resulted PCR fragment was isolated, cloned into pBluescript and sequenced. The isolated sequence encoding Trxh was identical to the published Trxh gene sequence
5 (Rivera-Madrid et al, (1993) Plant Physiol 102: 327-328). The pBluescript vector containing the Trxh gene is called pSBS2500.

For the isolation of the thioredoxin-reductase gene the following primers were synthesized:

GVR836: 5' GGCCAGCACACTACCATGAATGGTCTCGAAACTCAC 3' (SEQ ID
10 NO:3). The sequence identical to the 5' end of the thioredoxin-reductase gene as published (Jacquot et al, J Mol Biol. (1994) 235 (4):1357-63), is indicated in bold).

GVR837: 5' TTAAGCTTCAATCACTCTTACCTTGCTG 3' (SEQ ID NO:4).

A Polymerase Chain Reaction (PCR) was carried out using GVR836 and
15 GVR837 as primers and the cDNA library CD4-12 as a template. The resulted PCR fragment was isolated, cloned into pBluescript and sequenced. The pBluescript vector containing the thioredoxin-reductase gene is called pSBS2502.

A total of three clones were sequenced, the sequence of each of the
20 three clones were identical to each other. However, as depicted in Figure 1 this sequence indicated several nucleotide differences compared to the published thioredoxin-reductase gene sequence published (Jacquot et al, J Mol Biol. (1994) 235 (4):1357-63.). The complete coding sequence and its deduced amino acid sequence is shown in SEQ ID NO:10. As a result of the nucleotide differences
25 between the published sequence and the sequence isolated in Example 1, several amino acid changes are also predicted. A comparison of the deduced amino acid sequence of the published NADPH thioredoxin-reductase sequence thioredoxin-reductase (ATTHIREDB, Jacquot et al, J Mol Biol. (1994) 235 (4):1357-63.) with the sequence isolated in Example 1 (TR) is shown in Figure 3.

30

-81-

EXAMPLE 2**Construction of plant expression vectors.**

Expression vectors were constructed to allow for the seed specific over-expression of thioredoxin and NADPH thioredoxin-reductase in seeds. Vectors
5 were constructed to allow for over-expression in its natural subcellular location and for accumulation on oil bodies.

Construction of plant transformation vector pSBS2520.

The *Arabidopsis* thioredoxin h gene as described in example 1 was placed under the regulatory control of the phaseolin promoter and the phaseolin terminator
10 derived from the common bean *Phaseolus vulgaris* (Slightom et al (1983) Proc. Natl Acad Sc USA 80: 1897-1901; Sengupta-Gopalan *et al.*, (1985) PNAS USA 82: 3320-3324)). A gene splicing by overlap extension technique (Horton et al (1989) 15: 61-68) was used to fuse the phaseolin promoter to the Trxh gene. Standard molecular biology laboratory techniques (see eg: Sambrook *et al.*,
15 (1990) Molecular Cloning, 2nd ed. Cold Spring Harbor Press) were used to furnish the phaseolin promoter and terminator with Pst I and HindIII/KpnI sites respectively (see SEQ ID NO:14). Standard molecular biology laboratory techniques were also used to place the phaseolin terminator downstream from the Trxh gene. The PstI-phaseolin promoter- Trxh-phaseolin terminator-KpnI
20 insert sequence was cloned into the PstI-KpnI sites of pSBS3000 (pSBS3000 is a derivative from the *Agrobacterium* binary plasmid pPZP221 (Hajdukiewicz et al., 1994, Plant Molec. Biol. 25: 989-994). In pSBS3000, the CaMV35S promoter-gentamycin resistance gene-CAMV 35S terminator of pPZP221 was replaced with parsley ubiquitin promoter-phosphinothricin acetyl transferase
25 gene-parsley ubiquitin termination sequence to confer resistance to the herbicide glufosinate ammonium.) The resulting plasmid is called pSBS2520. The sequence of the phaseolin promoter-*Arabidopsis* Trxh-phaseolin terminator sequence is shown in SEQ ID NO:14.

Construction of plant transformation vector pSBS2510.

30 The 3' coding sequence of an *Arabidopsis* oleosin gene (van Rooijen et al (1992) Plant Mol. Biol. 18: 1177-1179) was altered to contain an NcoI site. The NcoI-HindIII fragment from vector pSBS2500 (Example 1) containing the Trxh was

-82-

ligated to the coding sequence of this *Arabidopsis* oleosin utilizing this NcoI restriction site. A gene splicing by overlap extension technique (Horton et al (1989) 15: 61-68) was used to fuse the phaseolin promoter (Slightom et al (1983) Proc. Natl Acad Sc USA 80: 1897-1901; Sengupta-Gopalan *et al.*,
5 (1985) PNAS USA 82: 3320-3324) containing a synthetic PstI site (see construction of pSBS2520) to the coding sequence of the *Arabidopsis* oleosin. Standard molecular biology laboratory techniques (see eg: Sambrook *et al.* (1990) Molecular Cloning, 2nd ed. Cold Spring Harbor Press) were again used to clone the HindIII KpnI fragment containing the phaseolin terminator (see
10 construction of pSBS2520) downstream of the Trxh gene. The PstI-phaseolin promoter- oleosin- Trxh-phaseolin terminator-KpnI insert sequence was cloned into the PstI-KpnI sites of pSBS3000. The resulting plasmid is called pSBS2510. The sequence of the phaseolin promoter-oleosin Trxh-phaseolin terminator sequence is shown in SEQ ID NO:16.

15 **Construction of plant transformation vector pSBS2521.**

This vector contains the same genetic elements as the insert of pSBS2510 except the Trxh gene is fused to the 5' end of the oleosin gene. The 3' oleosin coding sequence including its native stopcodon (van Rooijen et al (1992) Plant Mol. Biol. 18: 1177-1179) was furnished with a HindIII cloning site. Again a
20 gene splicing by overlap extension technique (Horton et al (1989) 15: 61-68) was used to fuse the phaseolin promoter to the Trxh gene and to fuse the Trxh gene to the oleosin sequence. Standard molecular biology laboratory techniques (see eg: Sambrook *et al.* (1990) Molecular Cloning, 2nd ed. Cold Spring Harbor Press) were again used to clone the HindIII KpnI fragment containing the
25 phaseolin terminator (see construction of pSBS2520) downstream of the oleosin gene. The PstI-phaseolin promoter- Trxh oleosin- phaseolin terminator-KpnI insert sequence was cloned into the PstI-KpnI sites of pSBS3000. The resulting plasmid is called pSBS2521. The sequence of the phaseolin promoter- Trxh oleosin -phaseolin terminator sequence is shown in SEQ ID NO:19.

30 **Construction of plant transformation vector pSBS2527.**

The *Arabidopsis* NADPH thioredoxin-reductase gene as described in example 1 was placed under the regulatory control of the phaseolin promoter and the

-83-

phaseolin terminator derived from the common bean *Phaseolus vulgaris* (Slightom et al (1983) Proc. Natl Acad Sc USA 80: 1897-1901; Sengupta-Gopalan *et al.*, (1985) PNAS USA 82: 3320-3324). A gene splicing by overlap extension technique (Horton et al (1989) 15: 61-68) was used to fuse the
5 phaseolin promoter to the thioredoxin-reductase gene. Standard molecular biology laboratory techniques (see eg: Sambrook *et al.* (1990) Molecular Cloning, 2nd ed. Cold Spring Harbor Press) were used to furnish the phaseolin promoter and terminator with PstI and HindIII/KpnI sites respectively (see SEQ ID NO:14). Standard molecular biology laboratory techniques were also used to place the
10 phaseolin terminator downstream from the thioredoxin-reductase gene. The PstI-phaseolin promoter-thioredoxin-reductase-phaseolin terminator-KpnI insert sequence was cloned into the PstI-KpnI sites of pSBS3000. The resulting plasmid is called pSBS2527. The sequence of the phaseolin promoter-*Arabidopsis* thioredoxin-reductase-phaseolin terminator sequence is shown in SEQ ID NO:22.

15 ***Construction of plant transformation vector pSBS2531.***

A gene splicing by overlap extension technique (Horton et al (1989) 15: 61-68) was used to fuse the phaseolin promoter (Slightom et al (1983) Proc. Natl Acad Sc USA 80: 1897-1901; Sengupta-Gopalan *et al.*, (1985) PNAS USA 82: 3320-3324) to the coding sequence of the *Arabidopsis* oleosin. The same gene
20 splicing technique was used to fuse the oleosin gene to the thioredoxin-reductase coding sequence. Standard molecular biology laboratory techniques (see eg: Sambrook *et al.* (1990) Molecular Cloning, 2nd ed. Cold Spring Harbor Press) were again used to clone the HindIII KpnI fragment containing the phaseolin downstream of the thioredoxin-reductase gene. The PstI-phaseolin
25 promoter- oleosin- thioredoxin-reductase -phaseolin terminator-KpnI insert sequence was cloned into the PstI-KpnI sites of pSBS3000. The resulting plasmid is called pSBS2531. The sequence of the phaseolin promoter-oleosin thioredoxin-reductase -phaseolin terminator sequence is shown in SEQ ID NO:24.

Construction of plant transformation vector pSBS2529

30 This vector contains the same genetic elements as the insert of pSBS2531 except the thioredoxin-reductase gene is fused to the 5' end of the oleosin gene. The 3' oleosin coding sequence including its native stopcodon (van Rooijen et al.

-84-

(1992) Plant Mol. Biol. 18: 1177-1179) was furnished with a HindIII cloning site. Again a gene splicing by overlap extension technique (Horton et al (1989) 15: 61-68) was used to fuse the phaseolin promoter to the thioredoxin-reductase gene and to fuse the thioredoxin-reductase gene to the oleosin sequence.

- 5 Standard molecular biology laboratory techniques (see eg: Sambrook *et al.* (1990) Molecular Cloning, 2nd ed. Cold Spring Harbor Press) were again used to clone the HindIII KpnI fragment containing the phaseolin terminator (see construction of pSBS2520) downstream of the oleosin gene. The PstI-phaseolin promoter- thioredoxin-reductase oleosin- phaseolin terminator-KpnI insert
- 10 sequence was cloned into the PstI-KpnI sites of pSBS3000. The resulting plasmid is called pSBS2529. The sequence of the phaseolin promoter- thioredoxin-reductase oleosin -phaseolin terminator sequence is shown in SEQ ID NO:27.

Construction of plant transformation vector pSBS2530.

- 15 A plant transformation was constructed containing the *Mycobacterium Leprae* thioredoxin-reductase /thioredoxin gene (*Mlep* TR/Trxh). A construct called pHIS/TR/Trxh (Wiele et al (1995) J Biol Chem 270:25604-25606) was obtained from the department of Immunohematology and Blood bank, Leiden University, The Netherlands and use as a template for PCR to generate pSBS2530. The
- 20 construction of pSBS2530 was identical to the construction of pSBS2531 except that the *Mlep* TR/Trxh gene was used instead of the *Arabidopsis* thioredoxin-reductase gene. A gene splicing by overlap extension technique (Horton et al (1989) 15: 61-68) was used to fuse the phaseolin promoter (Slightom et al (1983) Proc. Natl Acad Sc USA 80: 1897-1901; Sengupta-
- 25 Gopalan *et al.*, (1985) PNAS USA 82: 3320-3324) to the coding sequence of the *Arabidopsis* oleosin. The same gene splicing technique was used to fuse the oleosin gene to the *Mlep* TR/Trxh coding sequence. Standard molecular biology laboratory techniques (see eg: Sambrook *et al.* (1990) Molecular Cloning, 2nd ed. Cold Spring Harbor Press) were again used to clone the HindIII-KpnI fragment
- 30 containing the phaseolin downstream of the *Mlep* TR/Trxh gene. The PstI-phaseolin promoter- oleosin- *Mlep* TR/Trxh -phaseolin terminator-KpnI insert sequence was cloned into the PstI-KpnI sites of pSBS3000. The resulting

-85-

plasmid is called pSBS2530. The sequence of the phaseolin promoter-oleosin *Mlep* TR/Trxh -phaseolin terminator sequence is shown in SEQ ID NO:30.

Construction of plant transformation vector pSBS2542.

From initial activity assays (Figure 4), it was apparent that oil bodies expressing
 5 the oleosin-*M. lep* TR/Trxh fusion protein contained considerable reducing activity. It was anticipated that a similar oleosin fusion construct encoding the *Arabidopsis* thioredoxin-reductase and thioredoxin proteins would behave in an analogous manner. Molecular modeling was used to aid in the design of such a construct. Primers were designed (thioredoxin link-L: 5'-
 10 ACTGGAGATGTTGACTCGACGGATACTACGGATTGGTCGACGG
 CTATGGAAGAAGGACAAGTGATCGCCTGC-3'; (SEQ ID NO:5), and thioredoxin link-R:
 5'-ATCCGTCGAGTCAACATCTCCAGTTTCCTCGGTGGTCTCGTTAGCCTTCGAT
 CCAGCAATCTCTTGTAAGAATGCTCTGC-3'; (SEQ ID NO:6) to code for a
 15 synthetic linker peptide between the thioredoxin-reductase and thioredoxin proteins. These primers were used in conjunction with primers GVR 873 (5'-GTGGAAGCT TATGGAGATGGAG-3'; SEQ ID NO:7) and GVR834 (5'-GAAAGCTTAAGCCAAGTGTTTG-3'; SEQ ID NO:2) to amplify a region coding for a thioredoxin-reductase-linker region-thioredoxin utilizing a gene splicing by
 20 overlap extension technique (Horton et al (1989) 15:61-68). The thioredoxin-reductase-linker-thioredoxin encoding sequence was then cloned into a pre-existing pSBS3000 vector using standard molecular biology techniques (Sambrook et al (1990) Molecular Cloning 2nd Edition Cold Spring Harbour Press). The resulting plasmid was called pSBS2542. The sequence of the phaseolin
 25 promoter-oleosin-thioredoxin-reductase-linker-thioredoxin-phaseolin terminator region is shown in SEQ ID NO:33. An amino acid sequence comparison between this *Arabidopsis* thioredoxin-reductase-linker-thioredoxin and the *M. leprae* TR/Trxh protein is shown in Figure 12.

Plasmids pSBS2510, pSBS2520, pSBS2521, pSBS2527, pSBS2529,
 30 pSBS2530, pSBS2531 and pSBS2542 were electroporated into *Agrobacterium* strain EHA101. These *Agrobacterium* strains were used to transform *Arabidopsis*. *Arabidopsis* transformation was done essentially as described in

-86-

"*Arabidopsis* Protocols; Methods in molecular biology Vol 82. Edited by Martinez-Zapater JM and Salinas J. ISBN 0-89603-391-0 pg 259-266 (1998) except the putative transgenic plants were selected on agarose plates containing 80 μ M L-phosphinothricine, after they were transplanted to soil and allowed to set seed.

EXAMPLE 3

Polyacrylamide gelelectrophoresis and immunoblotting of transgenic seed extracts.

10 *Source of Arabidopsis thioredoxin, thioredoxin-reductase and oleosin antibodies.*

The *Arabidopsis* thioredoxin and thioredoxin-reductase genes were cloned in frame in bacterial expression vector pRSETB (Invitrogen) to allow for the overexpression of *Arabidopsis* thioredoxin and thioredoxin-reductase proteins.

15 These proteins were purified using standard protocols (see eg Invitrogen protocol) and used to raise antibodies in rabbits using standard biochemical techniques (See eg Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989). The *Arabidopsis* oleosin gene genes was cloned in frame in bacterial expression vector pRSETB (Invitrogen) to allow for the overexpression
20 *Arabidopsis* oleosin protein. This protein was purified using standard protocols (see eg Invitrogen protocol) and used to prepare mouse monoclonal antibodies using standard biochemical techniques (See eg Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989).

Preparation of total Arabidopsis seed extracts for PAGE. *Arabidopsis*
25 seeds were ground in approximately 20 volumes of 2% SDS, 50 mM Tris-Cl,, this extract was boiled, spun and the supernatant was prepared for polyacrylamide gelelectrophoresis (PAGE) using standard protocols.

Preparation of Arabidopsis oil-body-protein extracts.

Arabidopsis seeds were ground in approximately 20 volumes of water and spun
30 in a microfuge. The oil bodies were recovered and washed sequentially with approximately 20 volumes of water, a high stringency wash buffer, containing 8M urea and 100 mM sodiumcarbonate and water. After this last wash the

-87-

oil bodies are prepared for poly acrylamide gelelectrophoresis (PAGE) using standard protocols.

Analysis of seed and oil body extracts from plants transformed with pSBS2510

- 5 Total seed and oil body protein extracts from plants transformed with pSBS2510 were loaded onto polyacrylamide gels and either stained with coomassie brilliant blue or electroblotted onto PVDF membranes. The membranes were challenged with a polyclonal antibody raised against *Arabidopsis* thioredoxin, or a monoclonal antibody raised against the *Arabidopsis* 18.5 kDa
- 10 oleosin and visualized using alkaline phosphatase. Expression of the oleosin-thioredoxin results in an additional band of 31.2 kDa. The results indicate that the thioredoxin antibodies are immunologically reactive with a band of the right predicted molecular weight (31.2 kDa), and the oleosin antibodies are also immunologically reactive with a band of the right predicted molecular weight for
- 15 the fusion protein (31.2 kDa) in addition to a band corresponding to the native *Arabidopsis* oleosin (18.5 kDa). This indicates that oleosin-thioredoxin is expressed in *Arabidopsis* seeds and is correctly targeted to oil bodies.

Analysis of seed and oil body extracts from plants transformed with pSBS2521

- 20 Total seed and oil body protein extracts from plants transformed with pSBS25121 were loaded onto polyacrylamide gels and either stained with Coomassie brilliant blue or electroblotted onto PVDF membranes. The membranes were challenged with a polyclonal antibody raised against *Arabidopsis* thioredoxin, or a monoclonal antibody raised against the *Arabidopsis*
- 25 18.5 kDa oleosin and visualized using alkaline phosphatase. Expression of the thioredoxin-oleosin results in an additional band of 31.2 kDa. The results indicate that the thioredoxin antibodies are immunologically reactive with a band of the right predicted molecular weight (31.2 kDa), and the oleosin antibodies are also immunologically reactive with a band of the right predicted molecular
- 30 weight for the fusion protein (31.2 kDa) in addition to a band corresponding to the native *Arabidopsis* oleosin (18.5 kDa). This indicates that thioredoxin-oleosin is expressed in *Arabidopsis* seeds and is correctly targeted to oil bodies.

-88-

Analysis of seed extracts from plants transformed with pSBS2520 Total seed extracts from plants transformed with pSBS2520 were loaded onto polyacrylamide gels and either stained with Coomassie brilliant blue or electroblotted onto PVDF membranes. The membranes were challenged with a polyclonal antibody raised against *Arabidopsis* thioredoxin and visualized using alkaline phosphatase. The results indicated that the thioredoxin antibodies are immunologically reactive with a band of approximately the right predicted molecular weight (12 kDa). Untransformed seeds do not show a detectable thioredoxin band.

10 ***Analysis of seed and oil body extracts from plants transformed with pSBS2529***

Total seed and oilbody protein extracts from plants transformed with pSBS2529 were loaded onto polyacrylamide gels and electroblotted onto PVDF membranes. The membranes were challenged with a polyclonal antibody raised against *Arabidopsis* thioredoxin-reductase, or a monoclonal antibody raised against the *Arabidopsis* 18.5 kDa oleosin and visualized using alkaline phosphatase. Expression of the thioredoxin-reductase -oleosin results in an additional band of 53.8 kDa. The results indicate that the thioredoxin-reductase antibodies are immunologically reactive with a band of the right predicted molecular weight for the fusion protein (53.8 kDa), the oleosin antibodies are also immunologically reactive with a band of the right predicted molecular weight (53.8 kDa) in addition to a band corresponding to the native *Arabidopsis* oleosin (18.5 kDa). This indicates that thioredoxin-reductase-oleosin is expressed in *Arabidopsis* seeds.

25 ***Analysis of seed extracts from plants transformed with pSBS2527*** Total seed extracts from plants transformed with pSBS2527 were loaded onto polyacrylamide gels and electroblotted onto PVDF membranes. The membranes were challenged with a polyclonal antibody raised against *Arabidopsis* thioredoxin-reductase and visualized using alkaline phosphatase. The thioredoxin-reductase antibodies are immunologically reactive with a band of approximately the right predicted molecular weight for the (35.3 kDa). Untransformed seeds do not show a detectable thioredoxin band.

-89-

Analysis of seed extracts from plants transformed with pSBS2531

protein gel and immunoblot was prepared assaying the expression of oleosin-DMSR in *Arabidopsis* T2 seeds and correct targeting to *Arabidopsis* oil bodies. The expected molecular weight based on the deduced amino acid sequence is calculated to be 53,817 Da. In the oil body extract of the transgenic oleosin-thioredoxin-reductase sample an extra band of approximately 54 kDa was observed. This band was confirmed to be oleosin-thioredoxin-reductase by immunoblotting. From the polyacrylamide gel it was observed that the expression of the oleosin-Thioredoxin-reductase is about double compared to the expression of the major 18.5 kDa *Arabidopsis* oleosin. This represents approximately 2-4 % of total seed protein.

Analysis of seed extracts from plants transformed with pSBS2530

A protein gel and immunoblot was prepared assaying the expression of oleosin-*M.lep* TR/Trxh in *Arabidopsis* T2 seeds and the correct targeting to *Arabidopsis* oil bodies. The expected molecular weight based on the deduced amino acid sequence is calculated to be 67,550 Da. In the oil body extract of the transgenic oleosin-*M.lep* TR/Trxh sample an extra band of approximately 68 kDa was observed. This band was confirmed to be oleosin-*M.lep* TR/Trxh by immunoblotting. From the polyacrylamide gel it was observed that the expression of the oleosin-*M.lep* TR/Trxh is similar to the expression of the major 18.5 kDa *Arabidopsis* oleosin. This represents approximately 1-2 % of total seed protein.

Analysis of seed extracts from plants transformed with pSBS2542

Crude oil body extracts from pSBS2542 lines were prepared by grinding 100µg of seed in 1 mL of 100mM Tris buffer at pH 7.5. The samples were then centrifuged in order to isolate the oil body fraction. The oil body fraction was then loaded on an SDS polyacrylamide gel for expression analysis. A Coomassie stained gel revealed that the synthetic fusion accumulated to high levels in crude oil body extracts from 3 of the 4 lines tested. It was estimated that the fusion protein represented approximately 2-5% of total seed protein. Furthermore, western blots utilizing either anti-thioredoxin or anti-thioredoxin-reductase antibodies

-90-

confirmed that the over expressed 70 kDa protein was indeed oleosin-thioredoxin-reductase-linker-thioredoxin.

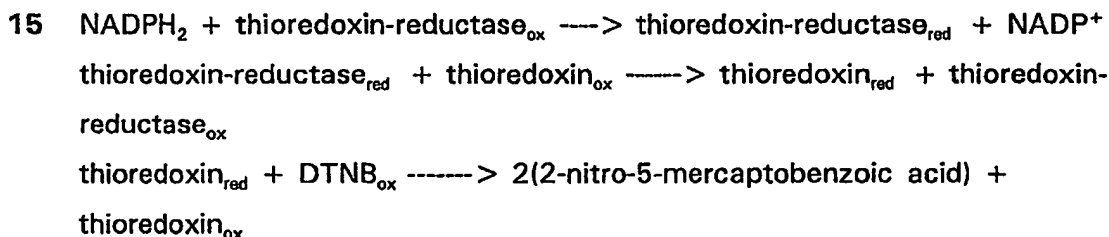
EXAMPLE 4

5 Biological activity of thioredoxin and thioredoxin-reductase transformants

Initial reduction assays:

DTNB assay

The activity of the thioredoxin and thioredoxin reductase was determined using a colorimetric DTNB [5,5'-dithiolbis (2-nitrobenzoic acid)] assay. The
 10 assay was performed in a 700 μ L reaction volume containing 100mM Tris-Cl pH 8.0, 5 mM EDTA, 200 μ M DTNB [5,5'-dithiolbis (2-nitrobenzoic acid)] and 200 μ M NADPH. If thioredoxin-reductase and thioredoxin are added, NADPH will reduce the thioredoxin-reductase, which will then reduce thioredoxin, which will, in turn, reduce DTNB (see equations below).



20 The formation of the yellow product was monitored by measuring the OD₄₁₂ in a spectrophotometer after a set period of time (usually 0.5-2 hours). The results of initial activity assays are shown in the bar graph in Figure 4 and described below.

Initially, 100 μ g of total seed proteins were added from each of the
 25 *Arabidopsis* transgenic lines, pSBS2520 (cytosolic thioredoxin) and pSBS2527 (cytosolic thioredoxin-reductase), which corresponds to approximately 1 μ g of cytosolic thioredoxin and thioredoxin-reductase used in the assay. In this case, the amount of DTNB reduced was comparable to the reduction caused by 1 μ g each of *E. coli* thioredoxin and thioredoxin-reductase. In these plant seed
 30 samples, background readings were very low when only one of the 2 extracts (either cytosolic thioredoxin or cytosolic thioredoxin-reductase; Figure 4, bars 3 and 6, respectively) was added to the reaction, along with wild type oil bodies.

-91-

Analysis with oil body fractions from transgenic seeds revealed that *Arabidopsis* thioredoxin and thioredoxin-reductase were substantially less active when fused to oleosins on oil bodies. Approximately 300 μ g of crude, unwashed oil-body-protein was used in the assay (which corresponds to 10-
5 30 μ g of thioredoxin-oleosin (pSBS 2521; Figure 4, bar 2), oleosin-thioredoxin (pSBS 2510, Figure 4, bar 1), thioredoxin-reductase-oleosin (pSBS 2529, Figure 4, bar 5), or oleosin-thioredoxin-reductase (pSBS 2531, Figure 4, bar 4). The oil-body-proteins were tested in conjunction with 100 μ g of total seed protein containing approximately 1 μ g of cytosolic thioredoxin (pSBS 2520) or
10 thioredoxin-reductase (pSBS 2527).

In such assays, pSBS2529 (thioredoxin-reductase-oleosin) and pSBS2531 (oleosin-thioredoxin-reductase) do contain reductase activity when combined with cytosolic thioredoxin from pSBS2520 (see Figure 4, bars 7 and 8, respectively). Experiments estimated that the reductase activity of oleosin-
15 thioredoxin-reductase was about 10-15% that of the cytosolic thioredoxin-reductase. The addition of tween at a final concentration of 0.4% could enhance this activity 2 or 3 fold. Interestingly, oleosin-thioredoxin-reductase (pSBS 2531) appears to be capable of reducing DTNB in the absence of added thioredoxin, although added thioredoxin causes significantly more DTNB
20 reduction (see Figure 4; compare bar 4 W.T. + oleosin-thioredoxin-reductase to bar 7 thioredoxin + oleosin-thioredoxin-reductase). Experiments with pSBS2521 (thioredoxin-oleosin) or pSBS2510 (oleosin-thioredoxin) combined with cytosolic thioredoxin-reductase from pSBS2527 (see Figure 4, bars 10 and 11, respectively) indicate that thioredoxin activity of these fusions is undetectable at
25 these concentrations.

Oil bodies from the transgenic *Arabidopsis* line, pSBS2530 (oleosin-*M.lep* TR/Trxh) contain significant thioredoxin/thioredoxin-reductase activity (see Figure 4, bar 12). One hundred micrograms of crude oil body protein for pSBS2530 was tested (corresponding to approximately 5 μ g of oleosin- *M.lep* TR/trxh
30 fusion) in the assay. Based on the assay, it was estimated that this fusion is about 25-40% as active as cytosolic *Arabidopsis* thioredoxin and thioredoxin-reductase (Figure 4, bar 9) when comparing specific activity.

-92-

Insulin reduction assay

The results from the DTNB assays were confirmed with insulin reduction assays. This assay contained insulin at a final concentration of 1mg/mL in 100mM KH_2PO_4 pH 7.0 + 5 mM EDTA. In the presence of NADPH (500 μ M),
5 thioredoxin, and thioredoxin-reductase, insulin is reduced and precipitates from the solution. Normally, insulin reduction is followed by measuring turbidity at OD 650. Alternatively, one can measure the conversion of NADPH₂ to NADP⁺ by monitoring the decrease in absorbance at 340 nm.

Both of the assays are difficult to measure when oil bodies are present,
10 due to interference with the spectrophotometer readings. However, qualitative data could be obtained by centrifuging the tubes after a set period of time, and determining if an insulin pellet was present (oil bodies float to the top, while the insulin precipitate pellets out). Alternatively, samples could be filtered after a set period of time, and the change in absorbance at 340 nm could be measured. As
15 mentioned previously, the results of the insulin reduction assays agreed with those of the DTNB assay, with the exception of the observation that pSBS2531 (oleosin-thioredoxin-reductase) only reduced insulin in the presence of free thioredoxin from pSBS2520.

Assays on seeds from Arabidopsis crosses that co-express oleosin-
20 ***thioredoxin and oleosin-thioredoxin-reductase.***

Based upon initial DTNB and insulin reduction assays, it was apparent that mixing oil bodies from oleosin<->thioredoxin and oleosin<->thioredoxin-reductase transgenic seeds resulted in very limited reducing activity (Note: the <-> indicates both configurations of oleosin fusions; ie. oleosin<->thioredoxin
25 would represent oleosin-thioredoxin and thioredoxin-oleosin fusions).

To determine whether having oleosin<->thioredoxin and oleosin<->thioredoxin-reductase proteins present on the same oil body would have a positive effect on the reducing activity of these proteins, crosses were set up to generate double transgenic *Arabidopsis* lines. The crosses are illustrated in
30 Table 2.

TABLE 2

	Male		Female	Confirmed double transgenic lines (PCR and Western Blot)
	oleo-thioredoxin	X	oleo-thioredoxin-reductase	4
5	oleo-thioredoxin	X	thioredoxin-reductase-oleo	1
	thioredoxin-oleo	X	oleo-thioredoxin-reductase	0
	thioredoxin-oleo	X	thioredoxin-reductase-oleo	4
	oleo-thioredoxin-reductase	X	oleo-thioredoxin	2
10	oleo-thioredoxin-reductase	X	thioredoxin-oleo	0
	thioredoxin-reductase-oleo	X	oleo-thioredoxin	7
15	thioredoxin-reductase-oleo	X	thioredoxin-oleo	0

Seeds from *Arabidopsis* crosses were germinated on PPT plates and the seedlings were transferred to soil after approximately 2 weeks. PCR experiments on DNA isolated from the seedlings identified a number of plants which contain both an oleosin<->thioredoxin and an oleosin<->thioredoxin-reductase gene construct within their genome.

Seeds were harvested from these plants for expression and activity assays. Western blots were carried out to confirm expression of both oleosin<->thioredoxin and oleosin<->thioredoxin-reductase in the lines. DTNB and insulin reduction assays were also performed to compare activity between single transgenic parent lines and the double transgenic offspring and results are summarized in Table 3. Table 3 summarizes DTNB reducing activity of various transgenic lines. The last 2 rows compare mixing oil bodies from single transgenic parent lines to using oil bodies from double transgenic offspring. Relative activity for the *E. coli* thioredoxin and thioredoxin mixture is set at 100 percent.

-94-

TABLE 3

	Source Material	Relative Activity (%)
	<i>E.coli</i> trx + NTR	100
5	<i>Arabidopsis</i> "free" thioredoxin + thioredoxin-reductase (pSBS2520 + pSBS2527)	100
	oleosin- <i>M. lep</i> TR/Trxh (pSBS2530)	~30
10	Oleosin<->thioredoxin-reductase + oleosin<->thioredoxin (mixing oil bodies from single-transgenic parents)	~3
15	Oleosin<->thioredoxin-reductase X oleosin<->thioredoxin (various double transgenic lines)	~50

Based on DTNB and insulin reduction assays, it is evident that double transgenic plants co-expressing oleosin<->thioredoxin and oleosin<->thioredoxin-reductase on the same, single oil body contained significantly more reducing activity compared to mixing oil bodies from single transgenic oleosin<->thioredoxin and oleosin<->thioredoxin-reductase lines. It was additionally apparent that oil body extracts from co-expressing lines contained more reducing activity compared to line pSBS2530 (oleosin-*M. lep* TR/Trxh), which was previously identified as the line containing the highest reducing activity from oil bodies.

These results suggest that the creation of double transgenic lines (either through crossing or by co-transforming 2 expression constructs into plants) may represent one means by which we could solve our initial problem of not being able to generate reducing activity by mixing oil bodies from oleosin<->thioredoxin and oleosin<->thioredoxin-reductase single transgenic lines.

Assays on seeds from *Arabidopsis* pSBS2542 transgenic lines that express oleosin-thioredoxin-reductase-linker-thioredoxin.

Oil body extracts from four pSBS2542 lines were tested for reducing activity in DTNB and insulin reduction assays, using standard protocols described previously. Again, oil body extracts containing the oleosin-thioredoxin-

-95-

reductase-linker-thioredoxin protein possessed significant reducing activity. Based on such assays, it was revealed that the oleosin-thioredoxin-reductase-linker-thioredoxin synthetic fusion protein was more active than the oleosin-*M. lep* TR/Trxh fusion. Furthermore, oil bodies containing the oleosin-thioredoxin-reductase-linker-thioredoxin protein appeared to have more reducing activity compared to oil bodies from double transgenic lines that co-expressed oleosin<->thioredoxin and oleosin<->thioredoxin-reductase. The results comparing reducing activity for the various thioredoxin-reductase/thioredoxin constructs is summarized in Table 4. Table 4 summarizes DTNB reducing activity of various transgenic lines. The pSBS2542 line expressing oleosin-thioredoxin-reductase-linker-thioredoxin contains significant reducing activity, comparable to the "free" forms of *Arabidopsis* thioredoxin and thioredoxin-reductase and the equivalent *E. coli* proteins. Relative activity for the *E. coli* thioredoxin and thioredoxin mixture is set at 100 percent.

15

TABLE 4

	Source Material	Relative Activity (%)
	<i>E. coli</i> trx + NTR	100
20	<i>Arabidopsis</i> "free" thioredoxin + thioredoxin-reductase (pSBS2520 + pSBS2527)	100
	oleosin- <i>M. lep</i> TR/Trxh (pSBS2530)	~30
25	Oleosin<->thioredoxin-reductase + oleosin<->thioredoxin (mixing oil bodies from single-transgenic parents)	~3
	Oleosin<->thioredoxin-reductase X oleosin<->thioredoxin (various double transgenic lines)	~50
30	Oleosin-thioredoxin-reductase-linker-thioredoxin (pSBS2542)	~75-100

-96-

Reduction assays comparing the utilization of NADH vs. NADPH as a cofactor (electron donor) for the thioredoxin-reductase/thioredoxin system.

DTNB and insulin reduction assays were conducted as described previously, except that NADH was substituted for NADPH as an electron donor in the system utilizing *E. coli* thioredoxin-reductase and thioredoxin. Thus, a comparison was conducted of the utilization of NADH versus NADPH as a cofactor for the *E. coli* thioredoxin-reductase/ thioredoxin system. For the DTNB assay, the reaction mixture consisted of 400 μ M DTNB, 10 μ g/mL *E. coli* thioredoxin, and 10 μ g/mL *E. coli* thioredoxin-reductase in 100mM Tris-Cl buffer pH 8.0. Either NADH or NADPH was then added to the DTNB reaction as follows:

- Reaction A. 200 μ M NADPH (Sigma)
- Reaction B. 800 μ M NADH (Sigma)
- Reaction C. 800 μ M NADH (Roche)
- 15 Reaction D. (-) cofactor
- Reaction E. 800 μ M NADH (no TR or Trxh).

For the insulin reduction assay, the reaction mixture consisted of 1 mg/mL bovine pancreatic insulin, 20 μ g/mL *E. coli* thioredoxin, and 20 μ g/mL *E. coli* thioredoxin-reductase in 100mM potassium phosphate buffer at pH 7.0. Either NADH or NADPH was then added to the reaction as follows:

- Reaction A. 800 μ M NADPH (Sigma)
- Reaction B. 800 μ M NADH (Sigma)
- Reaction C. 800 μ M NADH (Roche)
- 25 Reaction D. (-) cofactor
- Reaction E. 2 mM NADH (no TR or Trxh).

The results indicate that NADH, purchased from either Sigma or Roche, could act as an electron donor in both the DTNB and insulin reduction assays. However, the rate of reduction was lower than the rate observed with NADPH as a cofactor. It was estimated that the rate of insulin reduction utilizing NADH as an electron donor was approximately 25-50% when compared to the

-97-

maximum rate using NADPH. Furthermore, it was estimated that the rate of DTNB reduction utilizing NADH as an electron donor was approximately 5-10% of the maximum rate using NADPH. Similar results were observed using the oleosin-thioredoxin-reductase-linker thioredoxin fusion protein on *Arabidopsis* oil
5 bodies instead of the *E. coli* thioredoxin-reductase and thioredoxin.

Example 5

Production of multimeric immunoglobulin protein in plant seed cells and capture on oil bodies using Protein A – oleosin fusion proteins.

10 **1 – Production of multimeric immunoglobulin protein in plant seed cells**

For expression of multimeric-protein-complexes containing multimeric-immunoglobulin-complexes, the cDNA sequences encoding individual light and heavy chains can be isolated from; 1) cell lines expressing a particular antibody, such as clonal B cell lines, or a hybridoma cell line, or 2) may be a recombinant
15 antibody, assembled by combining select light and heavy chain variable domains and available light and heavy chain constant domain sequences, respectively. Variable domains with specific binding properties may be isolated from screening populations of such sequences, usually in the form of a single-chain Fv phage display library.

20 Starting from known nucleic acid sequences and a source of light and heavy chains, the mature polypeptide coding sequences of each chain is isolated with a secretion signal sequence. The signal sequence can be the native antibody sequence or derived from a known secreted plant sequence (e.g. a PR sequence from *Arabidopsis* or tobacco). The addition of a plant secretion signal
25 sequence to both light and heavy chain mature coding sequences is carried out by standard molecular biology techniques. PCR fusion is used routinely to make such modifications. Secretion signal sequences are included to target the light and heavy immunoglobulin polypeptides for secretion from the cell and further assembly of the two chains into a multimeric-immunoglobulin-complex. For
30 expression in transgenic plant seeds, an expression cassette is assembled comprising: 1) a regulatory promoter sequence to provide expression in plant seeds, 2) the secretion signal – light chain sequence, and 3) a regulatory

-98-

sequence to terminate transcription. A second expression cassette is assembled comprising: 1) a regulatory promoter sequence to provide expression in plant seeds, 2) the secretion signal – heavy chain sequence, and 3) a regulatory sequence to terminate transcription. Each of the antibody chain expression
5 cassettes is cloned individually into an Agrobacterium plant transformation vector or is combined into a single transformation vector with both expression cassettes. In both cases, the expression cassettes are cloned into plant transformation vectors, between the left and right delineating border sequences, and adjacent to a plant selectable marker cassette. Each plant transformation
10 vector is transformed into Agrobacterium. The resulting Agrobacterium strains are used to infect plant tissues. Transgenic plant material is regenerated and viable transgenic plants are selected. When individual transformation vectors are used, the transgenic plant lines that are produced, expressing either light or heavy chain sequences, are crossed to generate a single plant line expressing
15 both chains in the same plant cell. When a single transformation vector, containing both light and heavy expression cassettes, is used, the initial transgenic plant line produces both light and heavy chain sequences in the same plant cell.

2 – Production of transgenic oil bodies which display Protein A for the capture of
20 **immunoglobulins**

To capture and display immunoglobulin protein on oil bodies, oil bodies are engineered to display an immunoglobulin binding protein. In this example, the well-known antibody-binding domains from Protein A are used. Based on the known sequence for Protein A from *Staphylococcus aureus*, PCR primers are
25 designed to isolate the five consecutive Ig-binding domains from the bacterial Protein A sequence. Primers are designed to allow cloning of the Protein A sequence as either an N-terminal or C-terminal fusion to an oleosin sequence for targeting to oil bodies. The sequence that encodes an in-frame translational fusion between Protein A and oleosin is cloned into a plant expression cassette
30 for seed-specific expression. The final cassette consists of a regulatory promoter sequence that provides expression in seeds, the Protein A – oleosin fusion sequence, and a regulatory sequence to terminate transcription. The

Protein A - oleosin expression cassette is cloned into a plant transformation vector compatible with Agrobacterium – mediated plant transformation. The transformation vector comprises left and right border sequences flanking the Protein A – oleosin expression cassette and an adjacent plant selectable marker cassette. The Agrobacterium strain containing this vector is used to infect plant tissues and subsequent regeneration and selection from transgenic plant material to create transgenic plants.

3 – Capture and display of multimeric-immunoglobulins on oil bodies displaying Protein A

Having produced light and heavy chain multimeric immunoglobulin complexes in one transgenic plant line and the display of Protein A on oil bodies through the oil body targeting of a Protein A – oleosin fusion protein in a second plant line, at least two embodiments can be used to capture the immunoglobulin complex on the Protein A oil bodies. In the first embodiment, transgenic seed from both the immunoglobulin and the Protein A – oleosin expression lines is combined in an optimum ratio and then ground together such that the disrupted material from both seed lines would be combined in the same extract. The combined seed extracts are mixed and/or incubated under conditions that allow maximum recovery of the immunoglobulin by Protein A. The oil body fraction is separated using standard phase separation techniques (e.g. centrifugation). The recovered oil body fraction contains both native oil bodies, from the immunoglobulin expression line, and transgenic Protein A oil bodies from the Protein A – oleosin expression line.

In a second embodiment, the plant lines expressing the immunoglobulin complex and the Protein A – oleosin fusion are crossed and individual plant lines expressing both components are identified and propagated. In this approach, the immunoglobulin complex and the Protein A – oleosin fusion are produced in different cellular compartments of the same plant seed cell. Seed from the double transgenic line is ground to disrupt the cellular material and mix the contents of all cellular compartments, including combining the immunoglobulin in the extracellular compartment and the Protein A – oleosin on the oil body in the cytosolic compartment. The material is mixed and/or incubated under conditions

-100-

to allow maximum recovery of the immunoglobulin by Protein A, and the oil body fraction is separated by phase separation techniques. The recovered oil body fraction contains the displayed Protein A and the capture immunoglobulin complex.

5

Example 6

Production of assembled multimeric-immunoglobulin-complexes as fusions with oil body targeting domains.

Individual polypeptides are produced as a fusion protein with oil body targeting sequences (e.g. oleosin) for display on oil bodies. It has been found
10 that the individual subunits of naturally associating heterodimeric proteins can be co-produced as individual oleosin fusions and still associate as an active heterodimer on the surface of the oil body. In this example, the heterodimer is the light and heavy chain subunits, or derived portions thereof, of an immunoglobulin complex.

15 **Production of an immunoglobulin Fab complex on oil bodies.**

The mature light chain sequence, lacking the secretion signal sequence, is attached as an in-frame N-terminal fusion to an oleosin sequence. This fusion sequence is assembled into a seed-specific expression cassette consisting of a seed-specific promoter sequence, the light chain – oleosin fusion sequence, and
20 a transcriptional terminator sequence. The expression cassette is inserted between the left and right border markers, adjacent to a plant selectable marker cassette, of a transformation vector. The transformation vector, in *Agrobacterium*, is used to infect plants and generate transgenic plants.

An equivalent construct for the heavy chain subunit, comprising the
25 variable and constant heavy chain domains, is also attached as an in-frame fusion to oleosin and assembled into an expression cassette for seed-specific expression. The expression cassette can be a part of a separate transformation vector for the generation of a separate transgenic line, or the heavy chain expression cassette can be combined together with the light chain cassette into
30 a single transformation vector. If light and heavy chain expression cassettes are transformed into plants on separate transformation vectors, the individual plant lines are crossed to create a single line expressing both heterodimer subunit –

-101-

oleosin fusions in the same plant cell. Seed from the double transgenic line, or a single transgenic line generated from the dual expression vector, is extracted to isolate oil bodies. The seed material is ground to release the cellular contents and oil bodies are isolated by phase separation. The targeting of both light and heavy chain sequence to oil bodies, as oleosin fusions, allows the association of the immunoglobulin complex on the surface of the oil body.

Similar configurations, using the entire heavy chain sequence in combination with the entire light chain sequence, or using the variable domains from both the light and heavy chain sequences, are constructed to assemble different types of heteromultimeric-immunoglobulin-complexes (e.g., heterodimers) on the surface of oil bodies.

The present invention should therefore not be seen as limited to the particular embodiments described herein, but rather, it should be understood that the present invention has wide applicability with respect to protein expression generally. Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended claims.

-102-

SUMMARY OF SEQUENCES

SEQ ID NOs:1-4 set forth primers which were synthesized for the isolation of the thioredoxin h (Trxh) and thioredoxin reductase genes from *Arabidopsis*, as described in Example 1.

- 5 SEQ ID NOs:5-7 set forth primers which were designed to code for a specific linker peptide between thioredoxin reductase and thioredoxin proteins, as described in Example 2.

- SEQ ID NOs:8, 10 and 11 set forth the nucleotide sequence and the deduced amino acid sequence of the NADPH thioredoxin reductase sequence
10 isolated herein as described in Example 1.

 SEQ ID NOs:9 and 11, respectively, set forth the nucleotide sequence of the published NADPH thioredoxin reductase sequence (ATTHIREDB) and the deduced amino acid sequence.

- SEQ ID NO:12 sets forth the deduced amino acid sequence of the
15 published NADPH thioredoxin reductase sequence.

 SEQ ID NO:13 sets forth the deduced amino acid sequence of the NADPH reductase sequence isolated in this report.

- SEQ ID NOs:14 and 15 set forth the nucleotide sequence of the phaseolin promoter-*Arabidopsis* Trxh-phaseolin terminator sequence described in Example
20 2, and the deduced amino acid sequence. The Trxh coding sequence and its deduced amino acid sequence is indicated. The phaseolin promoter corresponds to nucleotide 6-1554, and the phaseolin terminator corresponds to nucleotide sequence 1905-3124. The promoter was furnished with a PstI site (nt 1-6) and the terminator was furnished with a HindIII site (nt 1898-1903) and a KpnI site
25 (nt 3124-3129) to facilitate cloning.

- SEQ ID NOs:16, 17 and 18 set forth the nucleotide sequence of the phaseolin promoter-oleosin Trxh-phaseolin terminator sequence described in Example 2, and the deduced amino acid sequences. The oleosin-Trxh coding sequence and the deduced amino acid sequences are indicated in SEQ ID NO:16.
30 As in SEQ ID NO:14, the phaseolin promoter corresponds to nucleotide 6-1554. The sequence encoding oleosin corresponds to nt 1555-2313, the intron in this sequence (nt 1908-2147) is indicated in italics. The Trxh coding sequence

-103-

corresponds to nt 2314-2658. The phaseolin terminator corresponds to nucleotide sequence 2664-3884.

SEQ ID NO:19, 20 and 21 set forth the nucleotide sequence of the phaseolin promoter - Trxh oleosin-phaseolin terminator sequence as described in Example 2, and the deduced amino acid sequences. The Trxh oleosin- coding sequence and its deduced amino acid sequences are indicated in SEQ ID NO:19. As in SEQ ID NOs:14 and 16, the phaseolin promoter corresponds to nucleotide 6-1554. The Trxh coding sequence corresponds to nt 1555-1896. The sequence encoding oleosin corresponds to nt 1897-2658, the intron in this sequence (nt 2250-2489) is indicated in italics. The phaseolin terminator corresponds to nucleotide sequence 2664-3884.

SEQ ID NO:22 and 23 set forth the nucleotide sequence of the phaseolin promoter-thioredoxin-reductase-phaseolin terminator sequence as described in Example 2, and the deduced amino acid sequence. The thioredoxin-reductase coding sequence and its deduced amino acid sequence is indicated in SEQ ID NO:22. The phaseolin promoter corresponds to nucleotide 6-1554. The thioredoxin-reductase coding sequence corresponds to nt 1555-2556 and the deduced amino acid is set forth in SEQ ID NO:23. The phaseolin terminator corresponds to nucleotide sequence 2563-3782.

SEQ ID NOs:24, 25 and 26 show the nucleotide sequence of the phaseolin promoter-oleosin thioredoxin-reductase-phaseolin terminator sequence as described in Example 2, and the deduced amino acid sequences. The oleosin-thioredoxin-reductase coding sequence and its deduced amino acid sequence is indicated. The phaseolin promoter corresponds to nucleotide 6-1554. The sequence encoding oleosin corresponds to nt 1555-2313, the intron in this sequence (nt 1980-2147) is indicated in italics. The thioredoxin-reductase coding sequence corresponds to nt 2314-3315. The phaseolin terminator corresponds to nucleotide sequence 3321-4540.

SEQ ID NOs:27, 28 and 29 show the nucleotide sequence of the phaseolin promoter - thioredoxin-reductase oleosin - phaseolin terminator sequence as described in Example 2, and the deduced amino acid sequences. The thioredoxin-reductase coding sequence and its deduced amino acid

-104-

sequence is indicated. The phaseolin promoter corresponds to nucleotide 6-1554. The thioredoxin-reductase coding sequence corresponds to nt 1555-2553. The sequence encoding oleosin corresponds to nt 2554-3315, the intron in this sequence (nt 2751-3146) is indicated in italics. The phaseolin terminator
 5 corresponds to nucleotide sequence 3321-4540.

SEQ ID NO:30, 31 and 32 show the sequence of the phaseolin promoter -oleosin - *Mlep* thioredoxin-reductase/thioredoxin -phaseolin terminator sequence as described in Example 2, and the deduced amino acid sequences. The oleosin-*Mlep* thioredoxin-reductase/thioredoxin coding sequence and its deduced amino
 10 acid sequence is indicated. The phaseolin promoter corresponds to nucleotide 6-1554. The sequence encoding oleosin corresponds to nt 1555-2313, the intron in this sequence (nt) is indicated in italics. The *Mlep* thioredoxin-reductase/thioredoxin coding sequence corresponds to nt 2314-3690. The phaseolin terminator corresponds to nucleotide sequence 3698-4917.

15 SEQ ID NOs:33, 34 and 35 set forth the nucleotide sequence of the phaseolin promoter-oleosin-thioredoxin-reductase-linker-thioredoxin-phaseolin terminator region of pSBS2542, and the deduced amino acid sequences. The deduced amino acid sequence of oleosin-thioredoxin-reductase-linker-thioredoxin is also shown in SEQ ID NO:33. Amino acids representing oleosin are set forth
 20 at positions 1-173, those amino acids representing thioredoxin-reductase are set forth at positions 174-501, those amino acids representing the linker or spacer peptide are set forth at positions 501-524, and those representing thioredoxin are set forth at positions 525-636.

SEQ ID NOs:38 and 39 set forth the nucleotide sequence of Arabidopsis
 25 *Thaliana* Thioredoxin h (Trx h 1) and the encoded protein, respectively.

SEQ ID NOs:40 and 41 set forth the nucleotide sequence of Arabidopsis *Thaliana* Thioredoxin Reductase (NTR1) and the encoded protein, respectively.

SEQ ID NOs:42 and 43 set forth the nucleotide sequence of *E. Coli* Thioredoxin (TrxA) and the encoded protein, respectively.

30 SEQ ID NOs:44 and 45, set forth the nucleotide sequence of *E. Coli* Thioredoxin Reductase and the encoded protein, respectively.

-105-

SEQ ID NOs:46 and 47 set forth the nucleotide sequence of Human Thioredoxin and the encoded protein, respectively.

SEQ ID NOs:48 and 49, set forth the nucleotide sequence of Human Thioredoxin Reductase and the encoded protein, respectively.

5 SEQ ID NOs:50 and 51, respectively, set forth the nucleotide sequence of Mycobacterium leprae Thioredoxin-Thioredoxin Reductase and the encoded protein, respectively.

SEQ ID NOs:52-313 are described in Table 5.

TABLE 5

10

15

20

25

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
PLANT THIOREDOXINS	
Thioredoxin f-type	
52	(Q9XFH8) Thioredoxin F-type 1, chloroplast precursor (TRX-F1). - Arabidopsis thaliana (Mouse-ear cress)
53	(Q9XFH9) Thioredoxin F-type 2, chloroplast precursor (TRX-F2). {GENE: AT5G16400 OR MQK4.13} - Arabidopsis thaliana (Mouse-ear cress)
54	(O48897) Thioredoxin F-type, chloroplast precursor (TRX-F). {GENE: TRXF} - Brassica napus (Rape)
55	(O81332) Thioredoxin F-type, chloroplast precursor (TRX-F). - Mesembryanthemum crystallinum (Common ice plant)
56	(P29450) Thioredoxin F-type, chloroplast precursor (TRX-F). - Pisum sativum (Garden pea)
57	(P09856) Thioredoxin F-type, chloroplast precursor (TRX-F). - Spinacia oleracea (Spinach)
Thioredoxin m-type	
58	(P06544) Thioredoxin 1 (TRX-1) (Thioredoxin M). {GENE: TRXA} - Anabaena sp. (strain PCC 7119)
59	(O48737) Thioredoxin M-type 1, chloroplast precursor (TRX-M1). {GENE: AT1G03680 OR F21B7_7 OR F21B7.28} - Arabidopsis thaliana (Mouse-ear cress)

-106-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
60	(Q9SEU8) Thioredoxin M-type 2, chloroplast precursor (TRX-M2). {GENE: AT4G03520 OR F9H3.15 OR T5L23.1} - Arabidopsis thaliana (Mouse-ear cress)
61	(Q9SEU7) Thioredoxin M-type 3, chloroplast precursor (TRX-M3). {GENE: AT2G15570 OR F9O13.12} - Arabidopsis thaliana (Mouse-ear cress)
62	(Q9SEU6) Thioredoxin M-type 4, chloroplast precursor (TRX-M4). - Arabidopsis thaliana (Mouse-ear cress)
63	(Q9XGS0) Thioredoxin M-type, chloroplast precursor (TRX-M). - Brassica napus (Rape)
5 64	(P23400) Thioredoxin M-type, chloroplast precursor (TRX-M) (Thioredoxin CH2). {GENE: TRXM} - Chlamydomonas reinhardtii
65	(Q41864) Thioredoxin M-type, chloroplast precursor (TRX-M). {GENE: TRM1} - Zea mays (Maize)
66	(Q9ZP20) Thioredoxin M-type, chloroplast precursor (TRX-M). - Oryza sativa (Rice)
67	(P48384) Thioredoxin M-type, chloroplast precursor (TRX-M). - Pisum sativum (Garden pea)
68	(P07591) Thioredoxin M-type, chloroplast precursor (TRX-M). - Spinacia oleracea (Spinach)
10 69	(Q9ZP21) Thioredoxin M-type, chloroplast precursor (TRX-M). - Triticum aestivum (Wheat)
70	(P12243) Thioredoxin 1 (TRX-1) (Thioredoxin M). {GENE: TRXA OR TRXM} - Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2)
71	(P37395) Thioredoxin. {GENE: TRXA OR TRX} - Cyanidium caldarium [Chloroplast]
72	(O22022) Thioredoxin. {GENE: TRXA OR TRXM} - Cyanidioschyzon merolae [Chloroplast]
73	(P50338) Thioredoxin. {GENE: TRXA} - Griffithsia pacifica [Chloroplast]

-107-

SEQ. ID NO..	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
74	(P50254) Thioredoxin. {GENE: TRXA} - Porphyra yezoensis [Chloroplast]
75	(P51225) Thioredoxin. {GENE: TRXA} - Porphyra purpurea [Chloroplast]
Thioredoxin h-type	
76	(P29448) Thioredoxin H-type 1 (TRX-H-1). {GENE: TRX1 OR AT3G51030 OR F24M12.70} - Arabidopsis thaliana (Mouse-ear cress)
5 77	(P20857) Thioredoxin 2 (TRX-2). {GENE: TRXB} - Anabaena sp. (strain PCC 7120)
78	(Q42388) Thioredoxin H-type 1 (TRX-H-1) (Pollen coat protein). {GENE: THL-1 OR BOPC17} - Brassica napus (Rape), Brassica oleracea (Cauliflower)
79	(P29449) Thioredoxin H-type 1 (TRX-H1). - Nicotiana tabacum (Common tobacco)
80	(Q38879) Thioredoxin H-type 2 (TRX-H-2). {GENE: TRX2 OR AT5G39950 OR MYH19.14} - Arabidopsis thaliana (Mouse-ear cress)
81	(Q39362) Thioredoxin H-type 2 (TRX-H-2). {GENE: THL-2} - Brassica napus (Rape)
10 82	(Q07090) Thioredoxin H-type 2 (TRX-H2). - Nicotiana tabacum (Common tobacco)
83	(Q42403) Thioredoxin H-type 3 (TRX-H-3). {GENE: TRX3 OR AT5G42980 OR MBD2.18} - Arabidopsis thaliana (Mouse-ear cress)
84	(Q39239) Thioredoxin H-type 4 (TRX-H-4). {GENE: TRX4} - Arabidopsis thaliana (Mouse-ear cress)
85	(Q39241) Thioredoxin H-type 5 (TRX-H-5). {GENE: TRX5} - Arabidopsis thaliana (Mouse-ear cress)
86	(O64432) Thioredoxin H-type (TRX-H). {GENE: PEC-2} - Brassica rapa (Turnip)

-108-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
87	(P80028) Thioredoxin H-type (TRX-H) (Thioredoxin CH1). {GENE: TRXH} - Chlamydomonas reinhardtii
88	(Q96419) Thioredoxin H-type (TRX-H). - Fagopyrum esculentum (Common buckwheat)
89	(Q42443) Thioredoxin H-type (TRX-H) (Phloem sap 13 kDa protein-1). - Oryza sativa (Rice)
90	(O65049) Thioredoxin H-type (TRX-H). {GENE: SB09} - Picea mariana (Black spruce)
5 91	(Q43636) Thioredoxin H-type (TRX-H). - Ricinus communis (Castor bean)
92	(O64394) Thioredoxin H-type (TRX-H) (TrxTa). - Triticum aestivum (Wheat)
93	(P29429) Thioredoxin. - Emericella nidulans (Aspergillus nidulans)
VIRUSES, BACTERIA AND FUNGI THIOREDOXINS	
94	(P80579) Thioredoxin (TRX). {GENE: TRXA} - Alicyclobacillus acidocaldarius (Bacillus acidocaldarius)
10 95	(O28137) Thioredoxin. {GENE: AF2145} - Archaeoglobus fulgidus
96	(P14949) Thioredoxin (TRX). {GENE: TRXA OR TRX} - Bacillus subtilis
97	(P00276) Thioredoxin. {GENE: NRDC} - Bacteriophage T4
98	(O51088) Thioredoxin (TRX). {GENE: TRXA OR BB0061} - Borrelia burgdorferi (Lyme disease spirochete)
99	(P57653) Thioredoxin (TRX). {GENE: TRXA OR BU597} - Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum symbiotic bacterium)
15 100	(O51890) Thioredoxin (TRX). {GENE: TRXA} - Buchnera aphidicola (subsp. Schizaphis graminum)
101	(P10472) Thioredoxin (TRX). {GENE: TRXA} - Chlorobium limicola f.sp. thiosulfatophilum

-109-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
102	(Q9PJK3) Thioredoxin (TRX). {GENE: TRXA OR TC0826} - Chlamydia muridarum
103	(Q9Z7P5) Thioredoxin (TRX). {GENE: TRXA OR CPN0659 OR CP0088} - Chlamydia pneumoniae (Chlamydophila pneumoniae)
104	(P52227) Thioredoxin (TRX). {GENE: TRXA} - Chlamydia psittaci (Chlamydophila psittaci)
105	(O84544) Thioredoxin (TRX). {GENE: TRXA OR CT539} - Chlamydia trachomatis
5	106 (P00275) Thioredoxin C-1. - Corynebacterium nephridii
	107 (P07887) Thioredoxin C-2. - Corynebacterium nephridii
	108 (P52228) Thioredoxin C-3. - Corynebacterium nephridii
	109 (P09857) Thioredoxin (TRX). {GENE: TRXA} - Chromatium vinosum
	110 (P21609) Thioredoxin (TRX). {GENE: TRXA} - Clostridium litoreale (Bacterium W6)
10	111 (P81108) Thioredoxin (TRX) (Fragment). {GENE: TRXA} - Clostridium sporogenes
	112 (P81109) Thioredoxin (TRX) (Fragment). {GENE: TRXA} - Clostridium sticklandii
	113 (Q9UW02) Thioredoxin (Allergen Cop c 2). - Coprinus comatus (Shaggy mane)
15	114 (P29445) Thioredoxin 1. {GENE: TRXA OR TRX1} - Dictyostelium discoideum (Slime mold)
	115 (P29446) Thioredoxin 2 (Fragment). {GENE: TRXB OR TRX2} - Dictyostelium discoideum (Slime mold)
	116 (P29447) Thioredoxin 3. {GENE: TRXC OR TRX3} - Dictyostelium discoideum (Slime mold)
	117 (P00274) Thioredoxin 1 (TRX1) (TRX). {GENE: TRXA OR TSNC OR FIPA OR B3781} - Escherichia coli, Salmonella typhimurium

-110-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
118	(P52232) Thioredoxin-like protein SLR0233. {GENE: SLR0233} - Synechocystis sp. (strain PCC 6803)
119	(P33636) Thioredoxin 2 (Trx2). {GENE: TRXC OR B2582 OR Z3867 OR ECS3448} - Escherichia coli, Escherichia coli O157:H7
120	(P21610) Thioredoxin (TRX). {GENE: TRXA} - Eubacterium acidaminophilum
121	(P43785) Thioredoxin (TRX). {GENE: TRXA OR TRXM OR HI0084} - Haemophilus influenzae
5 122	(P43787) Thioredoxin-like protein HI1115. {GENE: HI1115} - Haemophilus influenzae
123	(P56430) Thioredoxin (TRX). {GENE: TRXA OR HP0824 OR JHP0763} - Helicobacter pylori (Campylobacter pylori), Helicobacter pylori J99 (Campylobacter pylori J99)
124	(Q9S386) Thioredoxin (EC 1.6.4.5) {GENE:TRXA} - Listeria monocytogenes
125	(Q57755) Thioredoxin. {GENE: TRX OR MJ0307} - Methanococcus jannaschii
126	(P47370) Thioredoxin (TRX). {GENE: TRXA OR TRX OR MG124} - Mycoplasma genitalium
10 127	(P46843) Bifunctional thioredoxin-reductase/thioredoxin [Includes: Thioredoxin-reductase (EC 1.6.4.5) (TRXR); Thioredoxin]. {GENE: TRXB/A OR TRX OR ML2703} - Mycobacterium leprae
128	(P75512) Thioredoxin (TRX). {GENE: TRXA OR TRX OR MPN263 OR MP570} - Mycoplasma pneumoniae
129	(O30974) Thioredoxin (TRX). {GENE: TRXA} - Mycobacterium smegmatis
130	(P52229) Thioredoxin (TRX) (MPT46). {GENE: TRXA OR TRX OR TRXC OR RV3914 OR MT4033 OR MTV028.05} - Mycobacterium tuberculosis
131	(P42115) Thioredoxin. {GENE: TRX} - Neurospora crassa

-111-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
132	(P34723) Thioredoxin. {GENE: TRXA} - <i>Penicillium chrysogenum</i>
133	(Q9X2T1) Thioredoxin (TRX). {GENE: TRXA OR TRX OR PA5240} - <i>Pseudomonas aeruginosa</i>
134	(P10473) Thioredoxin (TRX). {GENE: TRXA} - <i>Rhodospirillum rubrum</i>
135	(P08058) Thioredoxin (TRX). {GENE: TRXA} - <i>Rhodobacter sphaeroides</i> (<i>Rhodopseudomonas sphaeroides</i>)
5	136 (Q9ZEE0) Thioredoxin (TRX). {GENE: TRXA OR RP002} - <i>Rickettsia prowazekii</i>
137	(P33791) Thioredoxin (TRX) (Fragment). {GENE: TRXA} - <i>Streptomyces aureofaciens</i>
138	(P52230) Thioredoxin (TRX). {GENE: TRXA OR SCH24.11C} - <i>Streptomyces coelicolor</i>
139	(Q05739) Thioredoxin (TRX). {GENE: TRXA} - <i>Streptomyces clavuligerus</i>
140	(P52231) Thioredoxin (TRX). {GENE: TRXA OR SLR0623} - <i>Synechocystis</i> sp. (strain PCC 6803)
10	141 (P73263) Thioredoxin-like protein SLR1139. {GENE: SLR1139} - <i>Synechocystis</i> sp. (strain PCC 6803)
142	(P52233) Thioredoxin (TRX). {GENE: TRXA} - <i>Thiobacillus ferrooxidans</i>
143	(P96132) Thioredoxin (TRX) (Fragment). {GENE: TRXA} - <i>Thiocapsa roseopersicina</i>
144	(P81110) Thioredoxin (TRX) (Fragment). {GENE: TRXA} - <i>Tissierella creatinophila</i>
145	(O83889) Thioredoxin (TRX). {GENE: TRXA OR TP0919} - <i>Treponema pallidum</i>

-112-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
ANIMAL THIOREDOXIN	
146	(O97680) Thioredoxin. {GENE: TXN} - Bos taurus (Bovine)
147	(Q95108) Thioredoxin, mitochondrial precursor (MT-TRX). {GENE: TXN2} - Bos taurus (Bovine)
148	(Q09433) Thioredoxin. {GENE: B0228.5} - Caenorhabditis elegans
5	149 (P99505) Thioredoxin (Fragment). {GENE: TXN} - Canis familiaris (Dog)
150	(P08629) Thioredoxin. {GENE: TXN} - Gallus gallus (Chicken)
151	(P47938) Thioredoxin (Deadhead protein). {GENE: DHD OR CG4193} - Drosophila melanogaster (Fruit fly)
152	(P10599) Thioredoxin (ATL-derived factor) (ADF) (Surface associated sulphhydryl protein) (SASP). {GENE: TXN OR TRDX OR TRX} - Homo sapiens (Human)
153	(Q99757) Thioredoxin, mitochondrial precursor (MT-TRX). {GENE: TXN2} - Homo sapiens (Human)
10	154 (P29451) Thioredoxin. {GENE: TXN} - Macaca mulatta (Rhesus macaque)
155	(P10639) Thioredoxin (ATL-derived factor) (ADF). {GENE: TXN} - Mus musculus (Mouse)
156	(P97493) Thioredoxin, mitochondrial precursor (MT-TRX). {GENE: TXN2} - Mus musculus (Mouse)
157	(P82460) Thioredoxin (Fragment). {GENE: TXN} - Sus scrofa (Pig)
158	(P08628) Thioredoxin. {GENE: TXN} - Oryctolagus cuniculus (Rabbit)
15	159 (P11232) Thioredoxin. {GENE: TXN} - Rattus norvegicus (Rat)
160	(P97615) Thioredoxin, mitochondrial precursor (MT-TRX). {GENE: TXN2 OR TRX2} - Rattus norvegicus (Rat)
161	(P50413) Thioredoxin. {GENE: TXN} - Ovis aries (Sheep)
PLANTS THIOREDOXIN-LIKE PROTEINS	

-113-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
162	(Q23166) THIOL-DISULFIDE INTERCHANGE LIKE PROTEIN (THIOREDOXIN-LIKE PROTEIN) {GENE:C7A10.160 OR AT4G37200 OR HCF164} - Arabidopsis thaliana (Mouse-ear cress)
163	(Q9C9Y6) Thioredoxin-like protein {GENE:F17O14.18} - Arabidopsis thaliana (Mouse-ear cress)
164	(Q9FYD5) Thioredoxin-like protein {GENE:F21E1_180} - Arabidopsis thaliana (Mouse-ear cress)
165	(Q38878) THIOREDOXIN-LIKE PROTEIN {GENE:TRX6 OR T7D17.3} - Arabidopsis thaliana (Mouse-ear cress)
5 166	(Q9LVI2) Thioredoxin-like protein - Arabidopsis thaliana (Mouse-ear cress)
167	(Q9SCN9) Thioredoxin-like protein {GENE:T4D2.150} - Arabidopsis thaliana (Mouse-ear cress)
168	(Q9SRD7) Thioredoxin-like protein, 49720-48645 {GENE:F28O16.13} - Arabidopsis thaliana (Mouse-ear cress)
169	(Q9SU84) THIOREDOXIN-LIKE PROTEIN {GENE:T16L4.180 OR AT4G29670} - Arabidopsis thaliana (Mouse-ear cress)
170	(Q9SWG6) Thioredoxin-like protein {GENE:TRX} - Hordeum bulbosum
10 171	(Q9SWG4) Thioredoxin-like protein {GENE:TRX} - Lolium perenne (Perennial ryegrass)
172	(Q9AS75) Thioredoxin-like protein {GENE:P0028E10.17} - Oryza sativa (Rice)
173	(O04002) CDSP32 protein (Chloroplast Drought-induced Stress Protein of 32kDa) - Solanum tuberosum (Potato)
174	(Q9SWG5) Thioredoxin-like protein {GENE:TRX} - Secale cereale (Rye)
175	(Q9SP36) Thioredoxin-like protein (Fragment) {GENE:TRX} - Secale cereale (Rye)
15 176	(Q9U515) Thioredoxin-like protein - Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm)

-114-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
VIRUSES, BACTERIA AND FUNGI THIOREDOXIN-LIKE PROTEINS	
177	(P43221) Thiol:disulfide interchange protein tlpA (Cytochrome c biogenesis protein tlpA). {GENE: TLPA} - Bradyrhizobium japonicum
178	(P43787) Thioredoxin-like protein HI1115. {GENE: HI1115} - Haemophilus influenzae
179	(Q9GUP7) Thioredoxin-like protein {GENE:TRXLP} - Leishmania major
5 180	(Q9UVH0) Thioredoxin-like protein - Mortierella alpina
181	(P95355) Thioredoxin-like protein - Neisseria gonorrhoeae
182	(Q98G37) Thioredoxin-like protein {GENE:MLL3505} - Rhizobium loti (Mesorhizobium loti)
183	(P36893) Thiol:disulfide interchange protein helX precursor (Cytochrome c biogenesis protein helX). {GENE: HELX} - Rhodobacter capsulatus (Rhodopseudomonas capsulata)
184	(P52232) Thioredoxin-like protein SLR0233. {GENE: SLR0233} - Synechocystis sp. (strain PCC 6803)
10 185	(P73263) Thioredoxin-like protein SLR1139. {GENE: SLR1139} - Synechocystis sp. (strain PCC 6803)
186	(Q9USR1) Thioredoxin-like protein {GENE:SPBC577.08C} - Schizosaccharomyces pombe (Fission yeast)
187	(Q9R788) Thioredoxin {GENE:TPTRX} - Treponema pallidum
ANIMALS THIOREDOXIN-LIKE PROTEINS	
188	(Q9UAV4) F46E10.9 PROTEIN (THIOREDOXIN-LIKE PROTEIN DPY-11) {GENE:F46E10.9 OR DPY-11} - Caenorhabditis elegans
15 189	(Q9N2K6) Thioredoxin-like protein (Y54E10A.3 protein) (Thioredoxin-like protein TXL) {GENE:TXL OR Y54E10A.3} - Caenorhabditis elegans
190	(Q9VRP3) THIOREDOXIN-LIKE PROTEIN TXL (CG5495 PROTEIN) {GENE:TXL OR CG5495} - Drosophila melanogaster (Fruit fly)

-115-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
191	(O43396) Thioredoxin-like protein (32 kDa thioredoxin-related protein). {GENE: TXNL OR TRP32 OR TXL} - Homo sapiens (Human)
192	(O76003) Thioredoxin-like protein - Homo sapiens (Human)
193	(Q9S753) THIOREDOXIN-LIKE PROTEIN {GENE:TRX} - Phalaris coerulescens
194	(O77404) TRYPAREDOXIN - Trypanosoma brucei brucei
5	PLANT THIOREDOXIN-REDUCTASES
195	(Q39243) Thioredoxin-reductase 1 (EC 1.6.4.5) (NADPH-dependent thioredoxin-reductase 1) (NTR 1). {GENE: NTR1 OR AT4G35460 OR F15J1.30} - Arabidopsis thaliana (Mouse-ear cress)
196	(Q39242) Thioredoxin-reductase 2 (EC 1.6.4.5) (NADPH-dependent thioredoxin-reductase 2) (NTR 2). {GENE: NTR2 OR AT2G17420 OR F5J6.18} - Arabidopsis thaliana (Mouse-ear cress)
10	VIRUSES, BACTERIA AND FUNGI THIOREDOXIN-REDUCTASES
197	(O66790) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR AQ_500} - Aquifex aeolicus
198	(P80880) Thioredoxin-reductase (EC 1.6.4.5) (TRXR) (General stress protein 35) (GSP35). {GENE: TRXB} - Bacillus subtilis
199	(P94284) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR BB0515} - Borrelia burgdorferi (Lyme disease spirochete)
200	(P57399) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR BU314} - Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum symbiotic bacterium)
201	(P81433) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB} - Buchnera aphidicola (subsp. Schizaphis graminum)
202	(Q9PKT7) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR TC0375} - Chlamydia muridarum

-116-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
203	(Q9Z8M4) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR CPN0314 OR CP0444} - Chlamydia pneumoniae (Chlamydophila pneumoniae)
204	(O84101) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR CT099} - Chlamydia trachomatis
205	(P52213) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB} - Clostridium litorale (Bacterium W6)
206	(P39916) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB} - Coxiella burnetii
5 207	(P09625) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR B0888 OR Z1232 OR ECS0973} - Escherichia coli, Escherichia coli O157:H7
208	(P50971) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB} - Eubacterium acidaminophilum
209	(P43788) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR HI1158} - Haemophilus influenzae
210	(Q9ZL18) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR JHP0764} - Helicobacter pylori J99 (Campylobacter pylori J99)
211	(P56431) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR HP0825} - Helicobacter pylori (Campylobacter pylori)
10 212	(O32823) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR LMO2478} - Listeria monocytogenes
213	(P47348) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR MG102} - Mycoplasma genitalium
214	(P46843) Bifunctional thioredoxin-reductase/thioredoxin [Includes: Thioredoxin-reductase (EC 1.6.4.5) (TRXR); Thioredoxin]. {GENE: TRXB/A OR TRX OR ML2703} - Mycobacterium leprae
215	(P75531) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR MPN240 OR MP591} - Mycoplasma pneumoniae
216	(O30973) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB} - Mycobacterium smegmatis

-117-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
217	(P52214) Thioredoxin-reductase (EC 1.6.4.5) (TRXR) (TR). {GENE: TRXB OR RV3913 OR MT4032 OR MTV028.04} - Mycobacterium tuberculosis
218	(P51978) Thioredoxin-reductase (EC 1.6.4.5). {GENE: CYS-9} - Neurospora crassa
219	(P43496) Thioredoxin-reductase (EC 1.6.4.5). {GENE: TRXB} - Penicillium chrysogenum
220	(Q9ZD97) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR RP445} - Rickettsia prowazekii
5 221	(Q92375) Thioredoxin-reductase (EC 1.6.4.5). {GENE: SPBC3F6.03} - Schizosaccharomyces pombe (Fission yeast)
222	(Q05741) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB} - Streptomyces clavuligerus
223	(P52215) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR SCH24.12} - Streptomyces coelicolor
224	(O83790) Thioredoxin-reductase (EC 1.6.4.5) (TRXR). {GENE: TRXB OR TP0814} - Treponema pallidum
225	(P80892) Thioredoxin-reductase (EC 1.6.4.5) (TRXR) (Fragment). {GENE: TRXB} - Vibrio fischeri
10 226	(P29509) Thioredoxin-reductase 1 (EC 1.6.4.5). {GENE: TRR1 OR YDR353W OR D9476.5} - Saccharomyces cerevisiae (Baker's yeast)
227	(P38816) Thioredoxin-reductase 2, mitochondrial precursor (EC 1.6.4.5). {GENE: TRR2 OR YHR106W} - Saccharomyces cerevisiae (Baker's yeast)
ANIMAL THIOREDOXIN-REDUCTASES	
228	(O62768) Thioredoxin-reductase (EC 1.6.4.5). {GENE: TXNRD1} - Bos taurus (Bovine)
229	(Q17745) Thioredoxin-reductase (EC 1.6.4.5). {GENE: C06G3.7} - Caenorhabditis elegans
15 230	(Q16881) Thioredoxin-reductase (EC 1.6.4.5). {GENE: TXNRD1} - Homo sapiens (Human)

-118-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
231	(Q25861) Thioredoxin-reductase (EC 1.6.4.5) (TrxR). {GENE: TR OR GR} - <i>Plasmodium falciparum</i> (isolate FCH-5)
Other thioredoxin-reductases	
PLANTS THIOREDOXIN-REDUCTASES	
232	(O22229) Thioredoxin-reductase {GENE:AT2G41680} - <i>Arabidopsis thaliana</i> (Mouse-ear cress)
5 233	(Q39951) NADPH thioredoxin-reductase (Fragment) - <i>Helianthus annuus</i> (Common sunflower)
VIRUSES, BACTERIA AND FUNGI THIOREDOXIN-REDUCTASES	
234	(O28718) THioredoxin-reductase (TRXB) {GENE:AF1554} - <i>Archaeoglobus fulgidus</i>
235	(Q9K703) Thioredoxin-reductase (NADPH) (EC 1.6.4.5) {GENE:TRXB OR BH3571} - <i>Bacillus halodurans</i>
236	(Q9K7F3) Thioredoxin-reductase {GENE:BH3408} - <i>Bacillus halodurans</i>
10 237	(Q9KCZ0) Thioredoxin-reductase {GENE:BH1429} - <i>Bacillus halodurans</i>
238	(Q9KCZ1) Thioredoxin-reductase {GENE:BH1428} - <i>Bacillus halodurans</i>
239	(Q9PIY1) Thioredoxin-reductase (EC 1.6.4.5) {GENE:TRXB OR CJO146} - <i>Campylobacter jejuni</i>
240	(Q9A4G3) Thioredoxin-reductase {GENE:CC2871} - <i>Caulobacter crescentus</i>
241	(Q97EM8) Thioredoxin-reductase {GENE:CAC3082} - <i>Clostridium acetobutylicum</i>
15 242	(Q97IU2) Thioredoxin-reductase {GENE:CAC1548} - <i>Clostridium acetobutylicum</i>
243	(Q9EV96) Thioredoxin-reductase {GENE:TRXB} - <i>Clostridium sticklandii</i>
244	(Q9RSY7) THioredoxin-reductase {GENE:DR1982} - <i>Deinococcus radiodurans</i>

-119-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
245	(O30739) Thioredoxin-reductase (Fragment) - Enterococcus faecalis (Streptococcus faecalis)
246	(O54535) Thioredoxin-reductase {GENE:TRXB OR TRXB1_2 OR VNG6452G OR TRXB1_1 OR VNG6074G} - Halobacterium sp. (strain NRC-1) [Plasmid pNRC100, and Plasmid pNRC200]
247	(P82854) Thioredoxin-reductase (EC 1.6.4.5) {GENE:TRXB2} - Halobacterium sp. (strain NRC-1)
248	(Q9HN08) Thioredoxin-reductase {GENE:TXRB3 OR VNG2301G} - Halobacterium sp. (strain NRC-1)
5 249	(O25779) Thioredoxin-reductase (TRXB) {GENE:HP1164} - Helicobacter pylori (Campylobacter pylori)
250	(O86255) Thioredoxin-reductase {GENE:TRXB} - Klebsiella oxytoca
251	(Q9AEV9) Thioredoxin-reductase (Fragment) {GENE:TRXB} - Lactococcus lactis (subsp. lactis) (Streptococcus lactis)
252	(Q9CF34) Thioredoxin-reductase (EC 1.6.4.5) {GENE:TRXB2} - Lactococcus lactis (subsp. lactis) (Streptococcus lactis)
253	(Q9CH02) Thioredoxin-reductase (EC 1.6.4.5) {GENE:TRXB1} - Lactococcus lactis (subsp. lactis) (Streptococcus lactis)
10 254	(Q9ZFC8) Thioredoxin-reductase (Fragment) {GENE:TRXB} - Lactococcus lactis
255	(O32822) Hypothetical 39.7 kDa protein (Fragment) - Listeria monocytogenes
256	(O26804) Thioredoxin-reductase {GENE:MTH708} - Methanothermobacter thermoautotrophicus
257	(P94397) Homologue of thioredoxin-reductase of Mycoplasma genitalium {GENE:YCGT} - Bacillus subtilis
258	(Q98PK9) Thioredoxin-reductase (EC 1.6.4.5) {GENE:MYPU_7130} - Mycoplasma pulmonis
15 259	(Q9JU23) Thioredoxin-reductase (EC 1.6.4.5) {GENE:TRXB OR NMA1538} - Neisseria meningitidis (serogroup A)

-120-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
260	(Q9JZ28) Thioredoxin-reductase {GENE:NMB1324} - <i>Neisseria meningitidis</i> (serogroup B)
261	(Q9IOM2) Thioredoxin-reductase 1 {GENE:TRXB1 OR PA2616} - <i>Pseudomonas aeruginosa</i>
262	(Q9I592) Thioredoxin-reductase 2 {GENE:TRXB2 OR PA0849} - <i>Pseudomonas aeruginosa</i>
263	(Q9V0Q8) Thioredoxin-reductase (TRXB) {GENE:TRXB OR PAB0500} - <i>Pyrococcus abyssi</i>
5 264	(Q9ZD33) Thioredoxin-reductase (TRXB2) {GENE:RP514} - <i>Rickettsia prowazekii</i>
265	(O54079) Thioredoxin-reductase (EC 1.6.4.5) {GENE:TRXB} - <i>Staphylococcus aureus</i>
266	(Q9RIS2) Thioredoxin-reductase {GENE:TRXB OR TRXB2} - <i>Streptomyces coelicolor</i>
267	(Q9K4L6) Thioredoxin-reductase {GENE:SC5F8.08C} - <i>Streptomyces coelicolor</i>
268	(Q97PY2) Thioredoxin-reductase {GENE:SP1458} - <i>Streptococcus pneumoniae</i>
10 269	(Q9A0B5) Thioredoxin-reductase {GENE:SPY0850} - <i>Streptococcus pyogenes</i>
270	(Q97V69) Thioredoxin-reductase (trxB-2) (EC 1.6.4.5) {GENE:TRXB-2} - <i>Sulfolobus solfataricus</i>
271	(Q97W27) Thioredoxin-reductase (trxB-3) (EC 1.6.4.5) {GENE:TRXB-3} - <i>Sulfolobus solfataricus</i>
272	(Q97WJ5) Thioredoxin-reductase (trxB-1) (EC 1.6.4.5) {GENE:TRXB-1} - <i>Sulfolobus solfataricus</i>
273	(Q98I59) Thioredoxin-reductase {GENE:MLL2552} - <i>Rhizobium loti</i> (<i>Mesorhizobium loti</i>)
15 274	(Q98M06) Thioredoxin-reductase {GENE:MLL0792} - <i>Rhizobium loti</i> (<i>Mesorhizobium loti</i>)

-121-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
275	(Q9UR80) 35 kDa THioredoxin-reductase HOMOLOG (FRAGMENT) {GENE:TRR1 AND YDR353W} - <i>Saccharomyces cerevisiae</i> (Baker's yeast)
276	(Q9ZE4) THIOREDOXIN {GENE:TRXA OR SA0992} - <i>Staphylococcus aureus</i> , <i>Staphylococcus aureus</i> subsp. <i>aureus</i> N315
277	(Q9S1H1) Thioredoxin-reductase (Fragment) {GENE:TRXB} - <i>Staphylococcus xylosus</i>
278	(Q9HJ14) Thioredoxin-reductase {GENE:TA0984} - <i>Thermoplasma acidophilum</i>
5 279	(Q9WZX3) THioredoxin-reductase {GENE:TM0869} - <i>Thermotoga maritima</i>
280	(Q979K8) Thioredoxin-reductase {GENE:TVG1183005} - <i>Thermoplasma volcanium</i>
281	(Q9PR71) Thioredoxin-reductase {GENE:TRXB OR UU074} - <i>Ureaplasma parvum</i> (<i>Ureaplasma urealyticum</i> biotype 1)
282	(Q9KSS4) Thioredoxin-reductase {GENE:VC1182} - <i>Vibrio cholerae</i>
283	(Q9PDD1) Thioredoxin-reductase {GENE:XF1448} - <i>Xylella fastidiosa</i>
10 284	(Q9X5F7) Thioredoxin-reductase {GENE:TRXB1} - <i>Zymomonas mobilis</i>
ANIMAL THIOREDOXIN-REDUCTASES	
285	(Q9GKW9) Thioredoxin-reductase 3 (Fragment) {GENE:TRXR3} - <i>Bos taurus</i> (Bovine)
286	(Q9N2I8) Thioredoxin-reductase (EC 1.6.4.5) - <i>Bos taurus</i> (Bovine)
287	(Q9N2K1) Thioredoxin-reductase homolog - <i>Caenorhabditis elegans</i>
15 288	(Q9NJH3) Thioredoxin-reductase - <i>Caenorhabditis elegans</i>

-122-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
289	(Q9VNT5) CG11401 PROTEIN (THioredoxin-reductase 2) {GENE:TRXR-2 OR CG11401} - Drosophila melanogaster (Fruit fly)
290	(O95840) Thioredoxin-reductase - Homo sapiens (Human)
291	(Q9UES8) Thioredoxin-reductase GRIM-12 - Homo sapiens (Human)
292	(Q9UH79) Thioredoxin-reductase {GENE:TR} - Homo sapiens (Human)
5 293	(Q9UQU8) Thioredoxin-reductase - Homo sapiens (Human)
294	(Q9NNW6) Thioredoxin-reductase TR2 (Fragment) - Homo sapiens (Human)
295	(Q9NNW7) Thioredoxin-reductase TR3 - Homo sapiens (Human)
296	(Q9P101) Thioredoxin-reductase 3 (Fragment) {GENE:TRXR3} - Homo sapiens (Human)
297	(Q9P2Y0) Thioredoxin-reductase II beta (EC 1.6.4.5) - Homo sapiens (Human)
10 298	(Q9H2Z5) Mitochondrial thioredoxin-reductase {GENE:TRXR2A} - Homo sapiens (Human)
299	(Q99475) KM-102-DERIVED REDUCTASE-LIKE FACTOR (THioredoxin-reductase) - Homo sapiens (Human)
300	(Q99P49) Thioredoxin-reductase 1 {GENE:TXNRD1} - Mus musculus (Mouse)
301	(Q9CSV5) Thioredoxin-reductase 1 (Fragment) {GENE:TXNRD1} - Mus musculus (Mouse)
302	(Q9CZE5) Thioredoxin-reductase 1 {GENE:TXNRD1} - Mus musculus (Mouse)
15 303	(Q9JHA7) Thioredoxin-reductase TR3 {GENE:TXNRD2 OR TR3} - Mus musculus (Mouse)
304	(Q9JLT4) Thioredoxin-reductase {GENE:TXNRD2 OR TRXR2} - Mus musculus (Mouse)

-123-

SEQ. ID NO.	SWISS PROTEIN IDENTIFIER (in parenthesis)
EXAMPLES OF REDOX PROTEINS	
305	(Q9JMH5) Thioredoxin-reductase 2 {GENE:TXNRD2 OR TXNRD2} - Mus musculus (Mouse)
306	(Q9JMH6) Thioredoxin-reductase 1 {GENE:TXNRD1 OR TXNRD1} - Mus musculus (Mouse)
307	(O89049) Thioredoxin-reductase - Rattus norvegicus (Rat)
308	(Q9JKZ3) Thioredoxin-reductase 1 (Fragment) - Rattus norvegicus (Rat)
309	(Q9JKZ4) Thioredoxin-reductase 1 - Rattus norvegicus (Rat)
310	(Q9JLE6) Thioredoxin-reductase (Fragment) - Rattus norvegicus (Rat)
311	(Q9R1I3) NADPH-dependent thioredoxin-reductase {GENE:TRR1} - Rattus norvegicus (Rat)
312	(Q9Z0J5) Thioredoxin-reductase precursor {GENE:TRXR2} - Rattus norvegicus (Rat)
313	(Q9MY8) Redox enzyme thioredoxin-reductase - Sus scrofa (Pig)

5

10

-124-

WHAT IS CLAIMED IS:

1. A method of producing an oil body associated with a recombinant multimeric-protein-complex, said method comprising:
 - (a) producing in a cell comprising oil bodies, a first recombinant polypeptide and
5 a second recombinant polypeptide wherein said first recombinant polypeptide is capable of associating with said second recombinant polypeptide to form said multimeric-protein-complex; and
 - (b) associating said multimeric-protein-complex with an oil body through an oil-body-targeting-protein capable of associating with said oil body and said first
10 recombinant polypeptide.
2. The method of claim 1 further comprising (c) isolating said oil bodies associated with said recombinant multimeric-protein-complex.
3. The method of claim 1 wherein said multimeric-protein-complex associates with oil bodies obtainable from said cell comprising oil bodies.
- 15 4. The method of claim 1 wherein said multimeric-protein-complex associates intracellularly with said oil bodies.
5. The method of claim 1 wherein said second recombinant polypeptide is associated with a second oil-body-targeting-protein capable of associating with an oil body and said second recombinant polypeptide.
- 20 6. The method of claim 5 wherein each of said oil-body-targeting-proteins is an oil-body-protein or an immunoglobulin.
7. The method of claim 6 wherein said oil-body-targeting-protein is an oleosin or caleosin.
8. The method of claim 1 wherein said oil-body-targeting-protein is an
25 oleosin or caleosin and said first recombinant polypeptide is fused to said oleosin or caleosin.
9. The method of claim 8 wherein said second recombinant polypeptide is fused to a second oleosin or second caleosin capable of associating with an oil body.
- 30 10. The method of claim 1 wherein said first and second recombinant polypeptides are produced as a multimeric-fusion-protein comprising said first and second recombinant polypeptide.

-125-

11. The method of claim 1, wherein said multimeric-protein-complex is a heteromultimeric-protein-complex.

12. The method of claim 11 wherein said heteromultimeric-protein-complex is an enzymatically active redox complex or an immunoglobulin.

5 13. The method of claim 1, wherein said first recombinant polypeptide is capable of associating with said second recombinant polypeptide in the cell.

14. The method of claim 1 wherein said first recombinant polypeptide is a thioredoxin and said second recombinant polypeptide is a thioredoxin-reductase.

10 15. The method of claim 14, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

16. The method of claim 14, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

15 17. The method of claim 1 wherein said cell is a plant cell.

18. The method of claim 1 wherein said cell is a safflower cell.

19. The method of claim 1 wherein said first recombinant polypeptide is an immunoglobulin-polypeptide-chain.

20 20. The method of claim 1 wherein said first recombinant polypeptide is an immunoglobulin light chain, or an immunologically active portion thereof, and said second recombinant polypeptide is an immunoglobulin heavy chain, or an immunologically active portion thereof.

21. The method of claim 19 wherein said oil-body targeting-protein comprises protein A, protein L or protein G.

25 22. A method of expressing a recombinant multimeric-protein-complex comprising a first and second recombinant polypeptide in a cell, said method comprising:

(a) introducing into a cell a first chimeric nucleic acid sequence comprising:

30 (i) a first nucleic acid sequence capable of regulating transcription in said cell operatively linked to;

(ii) a second nucleic acid sequence encoding a first recombinant polypeptide;

-126-

- (b) introducing into said cell a second chimeric nucleic acid sequence comprising:
- (i) a third nucleic acid sequence capable of regulating transcription in said cell operatively linked to;
 - (ii) a fourth nucleic acid sequence encoding a second recombinant polypeptide;
- (c) growing said cell under conditions to permit expression of said first and second recombinant polypeptide in a progeny cell comprising oil bodies wherein said first recombinant polypeptide and said second recombinant polypeptide are capable of forming a multimeric-protein-complex; and
- (d) associating said first recombinant polypeptide with an oil body through an oil-body-targeting-protein capable of associating with said oil body and said first recombinant polypeptide.
23. The method of claim 22 further comprising (e) isolating from said progeny cell, oil bodies comprising said multimeric-protein-complex.
24. The method of claim 22 wherein said multimeric-protein-complex associates with said oil bodies obtainable from said progeny cell comprising oil bodies.
25. The method of claim 22 wherein said oil bodies associate intracellularly with said multimeric-protein-complex.
26. The method of claim 22 wherein said second recombinant polypeptide is associated with a second oil-body-targeting-protein capable of associating with an oil body and said second recombinant polypeptide.
27. The method of claim 26 wherein each of said oil-body-targeting-proteins is selected from an oil-body-protein or an immunoglobulin.
28. The method of claim 27 wherein said oil-body-protein is an oleosin or caleosin.
29. The method of claim 28 wherein said first recombinant polypeptide is fused to said oleosin or caleosin.
30. The method of claim 29 wherein said second recombinant polypeptide is fused to a second oleosin or second caleosin capable of associating with an oil body.

-127-

31. The method of claim 22 wherein said first and second recombinant polypeptide are produced as a multimeric-fusion-protein comprising said first and second recombinant polypeptide.

32. The method of claim 22 wherein said first and second recombinant
5 polypeptide form a multimeric-protein-complex.

33. The method of claim 32, wherein said multimeric-protein-complex is a heteromultimeric-protein-complex.

34. The method of claim 32 wherein said heteromultimeric-protein-complex is an enzymatically active redox complex or an immunoglobulin.

10 35. The method of claim 22 wherein said first recombinant polypeptide and said second recombinant polypeptide are capable of forming a multimeric-protein-complex in said progeny cell.

36. The method of claim 22 wherein said first recombinant polypeptide is a thioredoxin and said second recombinant polypeptide is a thioredoxin-
15 reductase.

37. The method of claim 36, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

38. The method of claim 36, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10,
20 40, 44, 48, 50 and SEQ ID NOs:195-313.

39. The method of claim 22 wherein said first recombinant polypeptide is an immunoglobulin-polypeptide-chain.

40. The method of claim 22 wherein said first recombinant polypeptide is an immunoglobulin light chain, or an immunologically active portion thereof, and
25 said second recombinant polypeptide is an immunoglobulin heavy chain, or an immunologically active portion thereof.

41. The method of claim 39 wherein said oil-body targeting-protein comprises protein A, protein L or protein G.

30 42. The method of claim 22 wherein said cell is a plant cell.

43. The method of claim 42 wherein said plant cell is a safflower cell.

-128-

44. A method of producing in a plant a recombinant multimeric-protein-complex, said method comprising:

- (a) preparing a first plant comprising cells, said cells comprising oil bodies and a first recombinant polypeptide wherein said first recombinant polypeptide is
5 capable of associating with said oil bodies through an oil-body-targeting-protein;
(b) preparing a second plant comprising cells, said cells comprising oil bodies and a second recombinant polypeptide; and
(c) sexually crossing said first plant with said second plant to produce a progeny plant comprising cells, said cells comprising oil bodies, wherein said oil bodies
10 are capable of associating with said first recombinant polypeptide, and said first recombinant recombinant polypeptide is capable of associating with said second recombinant polypeptide to form said recombinant multimeric-protein-complex.

45. The method of claim 44 wherein said second recombinant polypeptide is capable of associating with oil bodies through an oil-body-targeting-protein in said second plant.
15

46. The method of claim 44 further comprising (d) isolating from said progeny plant oil bodies comprising said multimeric-protein-complex.

47. The method of claim 44 wherein said oil-body-targeting-protein is selected from an oil-body-protein or an immunoglobulin.

20 48. The method of claim 47 wherein said oil-body-protein is an oleosin or caleosin.

49. The method of claim 48 wherein said first recombinant polypeptide is fused to said oleosin or caleosin.

50. The method of claim 49 wherein said second recombinant polypeptide is fused to a second oleosin or second caleosin capable of associating with an oil body.
25

51. The method of claim 44 wherein said first and second recombinant polypeptide form a multimeric-protein-complex.

52. The method of claim 51, wherein said multimeric-protein-complex is a heteromultimeric-protein-complex.
30

53. The method of claim 52 wherein said heteromultimeric-protein-complex is an enzymatically active redox complex or an immunoglobulin.

-129-

54. The method of claim 44 wherein said first recombinant polypeptide is a thioredoxin and said second recombinant polypeptide is a thioredoxin-reductase.

55. The method of claim 54, wherein said thioredoxin is selected from
5 the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

56. The method of claim 54, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

57. The method of claim 44 wherein said first recombinant polypeptide is
10 an immunoglobulin-polypeptide-chain.

58. The method of claim 44 wherein said first recombinant polypeptide is an immunoglobulin light chain, or an immunologically active portion thereof, and said second recombinant polypeptide is an immunoglobulin heavy chain, or an immunologically active portion thereof.

59. The method of claim 57 wherein said oil-body targeting-protein
15 comprises protein A, protein L or protein G.

60. The method of claim 44 wherein said plant is safflower.

61. A chimeric nucleic acid sequence encoding a multimeric-fusion-protein, said nucleic acid comprising:

- 20 (a) a first nucleic acid sequence encoding an oil-body-targeting-protein operatively linked in reading frame to;
- (b) a second nucleic acid sequence encoding a first recombinant polypeptide; linked in reading frame to;
- (c) a third nucleic acid sequence encoding a second recombinant polypeptide,
25 wherein said first and second recombinant polypeptide are capable of forming a multimeric-protein-complex.

62. The nucleic acid of claim 61, wherein said oil-body-targeting-protein is selected from an oil-body-protein or an immunoglobulin.

63. The nucleic acid of claim 62, wherein said oil-body-protein is an
30 oleosin or caleosin.

64. The nucleic acid of claim 63, wherein said multimeric-protein-complex is a heteromultimeric-protein-complex.

-130-

65. The chimeric nucleic acid sequence of claim 61 wherein said first and second recombinant polypeptide form an enzymatically active heteromultimeric redox complex or an immunoglobulin.

66. The chimeric nucleic acid sequence of claim 65 wherein said first and second recombinant polypeptides are a thioredoxin and a thioredoxin-reductase.

67. The chimeric nucleic acid of claim 66, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

68. The chimeric nucleic acid of claim 66, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

69. The chimeric nucleic acid of claim 65 wherein said first recombinant polypeptide is an immunoglobulin-polypeptide-chain.

70. The chimeric nucleic acid of claim 65 wherein said first recombinant polypeptide is an immunoglobulin light chain, or an immunologically active portion thereof, and said second recombinant polypeptide is an immunoglobulin heavy chain, or an immunologically active portion thereof.

71. The chimeric nucleic acid of claim 69 wherein said oil-body targeting-protein comprises protein A, protein L or protein G.

72. The nucleic acid of claim 61, wherein positioned between said nucleic acid sequence encoding an oil-body-targeting-protein and said nucleic acid sequence encoding a first recombinant polypeptide is a linker nucleic acid sequence encoding an oil-body-surface-avoiding linker amino acid sequence.

73. The nucleic acid of claim 72, wherein said oil-body-surface-avoiding linker amino acid sequence is substantially negatively charged, or has a molecular weight of at least 35 kd.

74. The nucleic acid of claim 73, wherein the gene fusion further comprises a linker nucleic acid sequence encoding an amino acid sequence that is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is positioned between the oil-body-surface-avoiding linker amino acid sequence and said sequence encoding the first recombinant polypeptide.

-131-

75. A recombinant multimeric-fusion-protein comprising (i) an oil-body-targeting-protein, or fragment thereof, (ii) a first recombinant polypeptide and a (iii) second recombinant polypeptide, wherein said first and second recombinant polypeptides are capable of forming a multimeric-protein-complex.
- 5 76. The recombinant multimeric-fusion-protein of claim 75 wherein said oil-body-targeting-protein is selected from an oil-body-protein or an immunoglobulin.
77. The recombinant multimeric-fusion-protein of claim 76 wherein said oil-body-protein is an oleosin or a caleosin.
- 10 78. The recombinant multimeric-fusion-protein of claim 77, wherein said multimeric-fusion-protein is a heteromultimeric-fusion-protein.
79. The recombinant heteromultimeric-fusion-protein of claim 78 wherein said first and second recombinant polypeptide form an enzymatically active heteromultimeric redox complex or an immunoglobulin.
- 15 80. The recombinant fusion polypeptide of claim 79 wherein said first recombinant polypeptide is a thioredoxin and said second recombinant polypeptide is a thioredoxin-reductase.
81. The recombinant fusion polypeptide of claim 80, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50
20 and SEQ ID NOs:52-194.
82. The recombinant fusion polypeptide of claim 80, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.
83. The recombinant fusion polypeptide of claim 75, wherein positioned
25 between said oil-body-targeting-protein and said first recombinant polypeptide is an oil-body-surface-avoiding linker amino acid sequence.
84. The recombinant fusion polypeptide of claim 83, wherein said oil-body-surface-avoiding linker amino acid sequence is substantially negatively charged, or has a molecular weight of at least 35 kd.
- 30 85. The recombinant fusion polypeptide of claim 84, wherein the fusion polypeptide further comprises a linker amino acid sequence that is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is positioned

-132-

between the oil-body-surface-avoiding linker amino acid sequence and said first recombinant polypeptide.

86. Isolated oil bodies comprising a multimeric-protein-complex comprising (i) an oil-body-targeting-protein and (ii) a first recombinant
5 polypeptide, said oil bodies further comprising a second recombinant polypeptide, wherein said first and second recombinant polypeptide are capable of forming a multimeric-protein-complex.

87. Isolated oil bodies of claim 86 wherein said oil-body-targeting-protein is an oil-body-protein or an immunoglobulin.

10 88. Isolated oil bodies of claim 87 wherein said oil-body-protein is an oleosin or a caleosin.

89. Isolated oil bodies of claim 88 wherein said first recombinant polypeptide is fused to said oleosin or caleosin.

15 90. Isolated oil bodies of claim 86 wherein said first recombinant polypeptide is fused to said second recombinant polypeptide.

91. The isolated oil bodies of claim 90, wherein said multimeric-protein-complex is a heteromultimeric-protein-complex.

92. The isolated oil bodies of claim 91 wherein said heteromultimeric-protein-complex is an enzymatically active redox complex or an immunoglobulin.

20 93. Isolated oil bodies comprising
(a) a first fusion protein comprising a first oil-body-targeting-protein fused to a first recombinant polypeptide; and
(b) a second fusion protein comprising a second oil-body-targeting-protein fused to a second recombinant polypeptide,
25 wherein said first and second recombinant polypeptide are capable of forming a multimeric-protein-complex.

94. Isolated oil bodies of claim 93 wherein said first oil-body-targeting-protein is an oil-body-protein or an immunoglobulin.

30 95. Isolated oil bodies according claim 93 wherein said first oil-body-protein is an oleosin or a caleosin.

96. The isolated oil bodies of claim 93, wherein said multimeric-protein-complex is a heteromultimeric-protein-complex.

-133-

97. Isolated oil bodies of claim 93 wherein said first and second recombinant polypeptide form an enzymatically active heteromultimeric redox complex or an immunoglobulin.

98. Isolated oil bodies of claim 93 wherein said first recombinant
5 polypeptide is a thioredoxin and said second recombinant polypeptide is a thioredoxin-reductase.

99. The oil bodies of claim 98, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

100. The oil bodies of claim 98, wherein said thioredoxin-reductase is
10 selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

101. The oil bodies of claim 93 wherein said first recombinant polypeptide is an immunoglobulin-polypeptide-chain.

102. The oil bodies of claim 93 wherein said first recombinant
15 polypeptide is an immunoglobulin light chain, or an immunologically active portion thereof, and said second recombinant polypeptide is an immunoglobulin heavy chain, or an immunologically active portion thereof.

103. The oil bodies of claim 101 wherein said oil-body targeting-protein comprises protein A, protein L or protein G.

20 104. A cell comprising oil bodies and (i) an oil-body-targeting-protein, (ii) a first recombinant polypeptide and (iii) a second recombinant polypeptide wherein

(1) said first recombinant polypeptide is capable of associating with said oil-body-targeting-protein; and

25 (2) said first recombinant polypeptide capable of associating with said second recombinant polypeptide to form a multimeric-protein-complex.

105. The cell of claim 104 wherein said oil-body-targeting-protein is an oil-body-protein or an immunoglobulin.

106. The cell of claim 105 wherein said oil-body-protein is an oleosin or
30 caleosin.

-134-

107. The cell of claim 104 wherein said first recombinant polypeptide is fused to said second recombinant polypeptide so as to form a multimeric-fusion-protein.

108. The cell of claim 107 wherein said multimeric-fusion-protein is a
5 heteromultimeric-fusion-protein.

109. The cell of claim 104 wherein said first recombinant polypeptide is fused to said oil-body-targeting-protein.

110. The cell of claim 104 wherein said first recombinant polypeptide is fused to said first oil-body-targeting-protein and said second polypeptide is fused
10 to a second oil-body-targeting-protein.

111. The cell of claim 104 wherein said second recombinant polypeptide is capable of associating with a second oil-body-targeting-protein.

112. The cell of claim 104 wherein said first and second recombinant polypeptide form a heteromultimeric-protein-complex.

113. The cell of claim 104 wherein said heteromultimeric-protein-complex is an enzymatically active redox complex or an immunoglobulin.
15

114. The cell of claim 104 wherein said first polypeptide is a thioredoxin and said second polypeptide is a thioredoxin-reductase.

115. The cell of claim 114, wherein said thioredoxin is selected from the
20 group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

116. The cell of claim 114, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

117. The cell of claim 104 wherein said first recombinant polypeptide is
25 an immunoglobulin-polypeptide-chain.

118. The cell of claim 104 wherein said first recombinant polypeptide is an immunoglobulin light chain, or an immunologically active portion thereof, and said second recombinant polypeptide is an immunoglobulin heavy chain, or an immunologically active portion thereof.

119. The cell of claim 117 wherein said oil-body targeting-protein
30 comprises protein A, protein L or protein G.

120. The cell of claim 104 wherein said cell is obtained from a plant.

-135-

121. The cell of claim 104 wherein said cell is obtainable from a safflower plant.

122. A plant comprising cells of claim 104.

123. A safflower plant comprising cells of claim 104.

5 124. The method of claim 2 wherein said first recombinant polypeptide is a thioredoxin and said second recombinant polypeptide is a thioredoxin-reductase, said method further comprising (d) formulating the oil bodies for use in the preparation of a food product, personal care product or pharmaceutical composition.

10 125. The method of claim 124, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

126. The method of claim 124, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

15 127. The method of claim 124 wherein said formulating comprises the addition of NADP or NADPH.

128. The method of claim 124 wherein said food product is a milk or wheat based food product.

20 129. The method of claim 124 wherein said personal care product reduces the oxidative stress to the surface area of the human body or is used to lighten the skin.

25 130. The method of claim 124 wherein said pharmaceutical composition is used to treat chronic obstructive pulmonary disease (COPD), cataracts, diabetes, envenomation, bronchiopulmonary disease, malignancies, psoriasis, reperfusion injury, wound healing, sepsis, GI bleeding, intestinal bowel disease (IBD), ulcers, GERD (gastro esophageal reflux disease).

131. A composition comprising isolated oil bodies, thioredoxin and thioredoxin-reductase.

30 132. The composition of claim 131, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

-136-

133. The composition of claim 131, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

134. The composition of claim 131 further comprising NADP or NADPH.

5 135. A food product, personal care product or pharmaceutical composition comprising the composition of claim 131.

136. The food product, personal care product or pharmaceutical composition of claim 135 further comprising NADP or NADPH.

10 137. The food product of claim 135 wherein said food product is a milk based or wheat based food product.

138. The personal care product of claim 135 wherein said personal care product reduces the oxidative stress to the surface area of the human body or is used to lighten the skin.

15 139. The pharmaceutical composition of claim 135 wherein said pharmaceutical composition is used to treat chronic obstructive pulmonary disease, cataracts, psoriasis or reperfusion injury.

140. The multimeric-fusion-protein of claim 75, wherein said fusion-protein contains two or more polypeptide chains selected from the group of proteins set forth in Figure 5.

20 141. A method of reducing allergenicity of a food comprising the steps of:

 providing the isolated oil bodies of claim 78; and

 adding the isolated oil bodies to the food, whereby allergenicity of the food is reduced.

25 142. The method of claim 141, wherein the food is selected from the group consisting of wheat flour, wheat dough, milk, cheese, yogurt and ice cream.

143. The method of claim 141, further comprising providing NADH as a co-factor in the substantial absence of NADPH.

30 144. A method of treating or protecting a target against oxidative stress, comprising the steps of:

 providing the recombinant fusion polypeptide of claim 46; and

-137-

contacting the recombinant fusion polypeptide with a target, wherein the target is susceptible to oxidative stress, thereby treating or protecting against the stress.

145. The method of claim 144, wherein the target is selected from the group consisting of a molecule, a molecular complex, a cell, a tissue, and an organ.

146. A method for preparing an enzymatically active redox protein associated with oil bodies comprising:

- a) producing in a cell a redox fusion polypeptide comprising a first redox protein linked to a second redox protein;
- b) associating said redox fusion polypeptide with oil bodies through an oil-body-targeting-protein capable of associating with said redox fusion polypeptide and said oil bodies; and
- c) isolating said oil bodies associated with said redox fusion polypeptide.

147. The method of claim 146 wherein said oil-body-targeting-protein is an oil-body-protein or an immunoglobulin.

148. The method of claim 146 wherein said oil-body-protein is an oleosin or a caleosin.

149. The method of claim 146 wherein said first redox protein is a thioredoxin and said second redox protein is a thioredoxin-reductase.

150. The method of claim 149, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

151. The method of claim 149, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

152. The method of claim 146 wherein said cell is a plant cell.

153. The method of claim 146 wherein said cell is a safflower cell.

154. A method for preparing a redox protein associated with oil bodies comprising:

- a) introducing into a cell a chimeric nucleic acid sequence comprising:

-138-

- 1) a first nucleic acid sequence capable of regulating transcription in said cell operatively linked to;
- 2) a second nucleic acid sequence encoding a recombinant fusion polypeptide comprising (i) a nucleic acid sequence encoding a sufficient portion of an oil-body-protein to provide targeting of said recombinant fusion polypeptide to an oil body linked to (ii) a nucleic acid sequence encoding a redox fusion polypeptide comprising a first redox protein linked to a second redox protein operatively linked to;
- 3) a third nucleic acid sequence capable of terminating transcription in said cell;
- b) growing said cell under conditions to permit expression of said redox fusion polypeptide in a progeny cell comprising oil bodies; and
- c) isolating from said progeny cell said oil bodies comprising said redox fusion polypeptide.

155. The method of claim 154, wherein positioned between said nucleic acid sequence encoding a sufficient portion of an oil-body-protein and said nucleic acid sequence encoding a redox fusion polypeptide is a linker nucleic acid sequence encoding an oil-body-surface-avoiding linker amino acid sequence.

156. The method of claim 155, wherein said oil-body-surface-avoiding linker amino acid sequence is substantially negatively charged, or has a molecular weight of at least 35 kd.

157. The method of claim 156, wherein the gene fusion further comprises a linker nucleic acid sequence encoding an amino acid sequence that is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is positioned between the oil-body-surface-avoiding linker amino acid sequence and said nucleic acid sequence encoding a redox fusion polypeptide.

158. The method of claim 157, further comprising introducing an enzyme or chemical that cleaves said redox fusion polypeptide from said oil body, thereby obtaining isolated redox fusion polypeptide.

159. The method of claim 154 wherein said oil-body-protein is an oleosin or a caleosin.

-139-

160. The method of claim 154 wherein said first redox protein is a thioredoxin and said second redox protein is a thioredoxin-reductase.

161. The method of claim 160, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

5 162. The method of claim 160, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

163. The method of claim 154 wherein said cell is a plant cell.

10 164. The method of claim 154 wherein said thioredoxin and thioredoxin-reductase is obtained from *Arabidopsis*.

165. The method of claim 146 wherein the first redox protein is at least 5 times more active when produced as a redox fusion polypeptide as compared to the production of the first redox protein without the second redox protein.

166. The method of claim 146 further comprising:

15 d) formulating an emulsion of the oil bodies associated with the redox fusion polypeptide for use in the preparation of a product capable of treating oxidative stress in a target, a product capable of chemically reducing a target, pharmaceutical composition, a personal care product or a food product.

167. A chimeric nucleic acid comprising:

20 1) a first nucleic acid sequence capable of regulating transcription in a host cell operatively linked to;

2) a second nucleic acid sequence encoding a recombinant fusion polypeptide comprising (i) a nucleic acid sequence encoding a sufficient portion of an oil-body-protein to provide targeting of said recombinant fusion polypeptide to an oil body linked to (ii) a nucleic acid sequence encoding a redox fusion polypeptide comprising a first redox protein linked to a second redox protein operatively linked to;

25 3) a third nucleic acid sequence capable of terminating transcription in said cell.

30 168. The chimeric nucleic acid of claim 167 wherein said oil-body-protein is an oleosin or a caleosin.

-140-

169. The chimeric nucleic acid of claim 167 wherein said first redox protein is a thioredoxin and said second redox protein is a thioredoxin-reductase.

170. The chimeric nucleic acid of claim 169, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID
5 NOs:52-194.

171. The chimeric nucleic acid of claim 169, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

172. The chimeric nucleic acid of claim 167 wherein said cell is a plant
10 cell.

173. The chimeric nucleic acid of claim 167, wherein positioned between said nucleic acid sequence encoding a sufficient portion of an oil-body-protein and said nucleic acid sequence encoding a redox fusion polypeptide is a linker nucleic acid sequence encoding an oil-body-surface-avoiding linker amino acid
15 sequence.

174. The chimeric nucleic acid of claim 173, wherein said oil-body-surface-avoiding linker amino acid sequence is substantially negatively charged, or has a molecular weight of at least 35 kd.

175. The chimeric nucleic acid of claim 174, wherein the gene fusion
20 further comprises a linker nucleic acid sequence encoding an amino acid sequence that is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is positioned between the oil-body-surface-avoiding linker amino acid sequence and said nucleic acid sequence encoding a redox fusion polypeptide.

25 176. A transgenic plant comprising the chimeric nucleic acid sequence of claim 167.

177. The transgenic plant of claim 176, wherein said chimeric nucleic acid is contained within a plastid.

178. A safflower plant comprising the chimeric nucleic acid of anyone of
30 claim 167.

179. The safflower plant of claim 178, wherein said chimeric nucleic acid is contained within a plastid.

-141-

180. A plant seed comprising the chimeric nucleic acid of claim 167.

181. The plant seed of claim 180, wherein said chimeric nucleic acid is contained within a plastid.

182. A safflower seed comprising the chimeric nucleic acid of claim 168.

5 183. The safflower seed of claim 182, wherein said chimeric nucleic acid is contained within a plastid.

184. An oil body preparation obtained by the method of claim 146.

185. A food product comprising an oil body preparation of claim 184.

186. A composition comprising an oil body preparation of claim 184.

10 187. A personal care product comprising an oil body preparation of claim 184.

188. A product capable of treating oxidative stress in a target comprising an oil body preparation of claim 184.

15 189. A product capable of chemically reducing a target comprising an oil body preparation of claim 184.

190. A detergent composition comprising the product of claim 184.

191. A method of cleansing an item, comprising administering the product of claim 189 to said item under conditions that promote cleansing.

192. An emulsion formulation prepared by the method of claim 166.

20 193. A nucleic acid construct comprising a gene fusion, wherein the gene fusion comprises a first region encoding an oil-body-protein or an active fragment thereof, operably linked to a second region encoding at least one thioredoxin-related protein or an active fragment thereof.

25 194. The construct of claim 193, wherein the at least one thioredoxin-related protein is thioredoxin.

195. The construct of claim 194, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

30 196. The construct of claim 194, wherein the thioredoxin is derived from *Arabidopsis* or wheat.

197. The construct of claim 193, wherein the at least one thioredoxin-related protein is thioredoxin-reductase.

-142-

198. The construct of claim 197, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

199. The construct of claim 197, wherein the thioredoxin-reductase is
5 derived from *Arabidopsis* or wheat.

200. The construct of claim 197, wherein the thioredoxin-reductase is an NADPH-dependent thioredoxin-reductase.

201. The construct of claim 193, wherein the second region encodes a thioredoxin and thioredoxin-reductase.

10 202. The construct of claim 201, wherein the thioredoxin and thioredoxin-reductase is obtained from *Mycobacterium leprae*.

203. The construct of claim 201, wherein at least one thioredoxin-related protein is an engineered fusion protein.

15 204. The construct of claim 193; wherein the first region precedes, in a 5' to 3' direction, the second region.

205. The construct of claim 193, wherein the first region follows, in a 5' to 3' direction, the second region.

206. The construct of claim 193, wherein the gene fusion further comprises a third region encoding a second thioredoxin-related protein or an
20 active fragment thereof, operably linked to the first region, or to the second region, or to both.

207. The construct of claim 193, further comprising a seed-specific promoter operably linked to the gene fusion.

208. The construct of claim 207, wherein the promoter is a phaseolin
25 promoter.

209. The construct of claim 193, wherein at least one thioredoxin-related protein is derived from a plant species selected from the group consisting of *Arabidopsis* and wheat.

210. The construct of claim 193, wherein at least one thioredoxin-related
30 protein is derived from *E. coli*.

-143-

211. The construct of claim 193 further comprising a nucleic acid effective as a termination region in plant cells, operably linked to the gene fusion.

212. The construct of claim 193, wherein the gene fusion further
5 comprises a nucleic acid sequence encoding an oil-body-surface-avoiding linker amino acid sequence, wherein the linker amino acid sequence is positioned between the first region and the second region.

213. The construct of claim 212, wherein said oil-body-surface-avoiding linker amino acid sequence is substantially negatively charged, or has a
10 molecular weight of at least 35 kd.

214. The construct of claim 213, wherein the gene fusion further comprises a linker nucleic acid sequence encoding an amino acid sequence that is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is positioned between the oil-body-surface-avoiding linker amino acid sequence
15 and the second region.

215. The construct of claim 193, wherein a region of the gene fusion comprises a plurality of codons, each codon specifying a single amino acid, wherein at least one of the codons is modified from a naturally occurring codon within the region.

20 216. The construct of claim 215, wherein the modified codon specifies the same amino acid as the naturally occurring codon, and wherein the modified codon is modified according to a codon preference of a plant.

217. The construct of claim 215, wherein the modified codon specifies an amino acid that is different from the amino acid specified by the naturally
25 occurring codon.

218. A transgenic plant containing a nucleic acid construct comprising a gene fusion, wherein the gene fusion comprises a region encoding an oil-body-protein or an active fragment thereof, operably linked to a region encoding a first thioredoxin-related protein or an active fragment thereof.

30 219. The plant of claim 218, wherein the thioredoxin-related protein is thioredoxin.

-144-

220. The plant of claim 219, wherein said thioredoxin is selected from the group consisting of SEQ ID NOs:38, 42, 46, 50 and SEQ ID NOs:52-194.

221. The plant of claim 219, wherein the thioredoxin is derived from *Arabidopsis* or wheat.

5 222. The plant of claim 218, wherein the thioredoxin-related protein is thioredoxin-reductase.

223. The plant of claim 222, wherein said thioredoxin-reductase is selected from the group consisting of those set forth in SEQ ID NOs:8, 9, 10, 40, 44, 48, 50 and SEQ ID NOs:195-313.

10 224. The plant of claim 222, wherein the thioredoxin-reductase is an NADPH-dependent thioredoxin-reductase.

225. The plant of claim 218, wherein the construct is contained within a plastid.

226. The plant of claim 218, wherein the first thioredoxin-related protein
15 is thioredoxin and wherein said construct further comprises a region encoding a thioredoxin-reductase.

227. The plant of claim 226, wherein the thioredoxin and thioredoxin-reductase is obtained from *Mycobacterium leprae*.

228. The plant of claim 226, wherein the thioredoxin-related protein is
20 an engineered fusion protein.

229. The plant of claim 218, wherein the first region precedes, in a 5' to 3' direction, the second region.

230. The plant of claim 218, wherein the first region follows, in a 5' to 3' direction, the second region.

25 231. The plant of claim 218, wherein the gene fusion further comprises a third region encoding a second thioredoxin-related protein or an active fragment thereof, operably linked to the first region, or to the second region, or to both.

232. The plant of claim 218, further comprising a seed-specific promoter
30 operably linked to the gene fusion.

233. The plant of claim 232, wherein the promoter is a phaseolin promoter.

-145-

234. The plant of claim 218, wherein the thioredoxin-related protein is derived from a plant species selected from the group consisting of *Arabidopsis* and wheat.

235. The plant of claim 218, wherein the thioredoxin-related protein is
5 derived from *E. coli*.

236. The plant of claim 218 further comprising a nucleic acid effective as a termination region in plant cells, operably linked to the gene fusion.

237. The plant of claim 218, wherein the gene fusion further comprises a nucleic acid sequence encoding an oil-body-surface-avoiding linker amino acid
10 sequence, wherein the nucleic acid encoding the linker amino acid sequence is positioned between the region encoding an oil-body-protein and the region encoding a first thioredoxin-related protein.

238. The plant of claim 237, wherein said oil-body-surface-avoiding linker amino acid sequence is substantially negatively charged, or has a
15 molecular weight of at least 35 kd.

239. The plant of claim 238, wherein the gene fusion further comprises a linker nucleic acid sequence encoding an amino acid sequence that is specifically cleavable by an enzyme or a chemical, wherein the linker sequence is positioned between the oil-body-surface-avoiding linker amino acid sequence and
20 the region encoding a first thioredoxin-related protein.

240. The plant of claim 218, wherein a region of the gene fusion comprises a plurality of codons, each codon specifying a single amino acid, wherein at least one of the codons is modified from a naturally occurring codon within the region.

25 241. The plant of claim 240, wherein the modified codon specifies the same amino acid as the naturally occurring codon, and wherein the codon is modified according to a codon preference of a plant.

242. The plant of claim 240, wherein the modified codon specifies an amino acid that is different from the amino acid specified by the naturally
30 occurring codon.

-146-

243. The plant of claim 218, wherein the plant is selected from the group consisting of *Arabidopsis* and safflower.

244. A transgenic plant comprising a nucleic acid construct a seed-specific promoter operably linked to a gene fusion, wherein the gene fusion
5 comprises a region encoding an oil-body-protein or an active fragment thereof, operably linked to a region encoding a first thioredoxin-related protein or an active fragment thereof, wherein a fusion protein comprising activities of oleosin and the thioredoxin-related protein is produced in a seed of the plant.

245. The transgenic plant of claim 244, wherein the plant is selected
10 from the group consisting of *Arabidopsis* and safflower.

246. The transgenic plant of claim 244 wherein the promoter is a phaseolin promoter.

247. The seed of the plant of claim 244.

248. The seed of claim 247, comprising a thioredoxin-related protein in a
15 concentration of at least about 0.5% of total cellular seed protein.

249. An extract of the seed of claim 247, wherein the extract comprises an activity of a thioredoxin-related protein.

250. An oil body from the seed of claim 247.

251. Oil produced from the seed of claim 247.

252. A method of making a fusion protein comprising a thioredoxin-
20 related activity, the method comprising the steps of:

providing a transgenic plant comprising a nucleic acid construct
comprising a seed-specific promoter operably linked to a gene
fusion, wherein the gene fusion comprises a region encoding an
25 oil-body-protein or an active fragment thereof, operably linked to a region encoding a first thioredoxin-related protein or an active fragment thereof, the gene fusion encoding a fusion protein comprising a thioredoxin-related activity;

obtaining seeds from the plant; and
30 recovering the fusion protein by isolating oil bodies from the seeds.

253. The method of claim 252, further comprising the step of fractionating the oil bodies to achieve partial purification of the fusion protein.

-147-

254. Oil bodies in association with a fusion protein, obtained by the method of claim 252.

255. The method of claim 252 further comprising a step of cleaving the oil-body-protein from the thioredoxin-related protein after fractionation of the oil
5 bodies.

256. The method of claim 255, wherein the cleaving step comprises use of a protease.

257. The method of claim 255, wherein the cleaving step comprises chemical proteolysis.

10 258. A method of reducing allergenicity of a food comprising the steps of:

providing a preparation comprising oil bodies associated with a fusion protein, the fusion protein comprising an oil-body-protein or an active fragment thereof and a thioredoxin-related protein or an active
15 fragment thereof; and

adding the preparation to the food, whereby allergenicity of the food is reduced due to activity of the thioredoxin-related protein or fragment.

259. The method of claim 258, wherein the food is selected from the
20 group consisting of wheat flour, wheat dough, milk, cheese, yogurt and ice cream.

260. The method of claim 258, further comprising providing NADH as a co-factor in the substantial absence of NADPH.

261. A composition comprising a fusion protein, the fusion protein
25 comprising an oil-body-protein or an active fragment thereof and a thioredoxin-related protein or an active fragment thereof, in a pharmaceutically acceptable carrier.

262. The composition of claim 261, further comprising oil bodies in association with the fusion protein.

30 263. A cosmetic formulation comprising oil bodies associated with a fusion protein, the fusion protein comprising an oil-body-protein or an active

-148-

fragment thereof and a thioredoxin-related protein or an active fragment thereof, in a pharmaceutically acceptable carrier.

264. A method of treating or protecting a target against oxidative stress, comprising the steps of:

- 5 providing a preparation comprising a fusion protein, the fusion protein comprising an oil-body-protein or an active fragment thereof and a thioredoxin-related protein or an active fragment thereof; and
- contacting the preparation with a target, wherein the target is susceptible to oxidative stress, thereby treating or protecting against the
- 10 stress.

265. The method of claim 264, wherein the target is selected from the group consisting of a molecule, a molecular complex, a cell, a tissue, and an organ.

266. A nucleic acid construct comprising a gene fusion, wherein the
- 15 gene fusion comprises a first region encoding an oil-body-protein or an active fragment thereof, operably linked to a second region encoding at least one polypeptide or an active fragment thereof, and an oil-body-surface-avoiding linker in frame between the first and second region polypeptides.

FIGURE 1

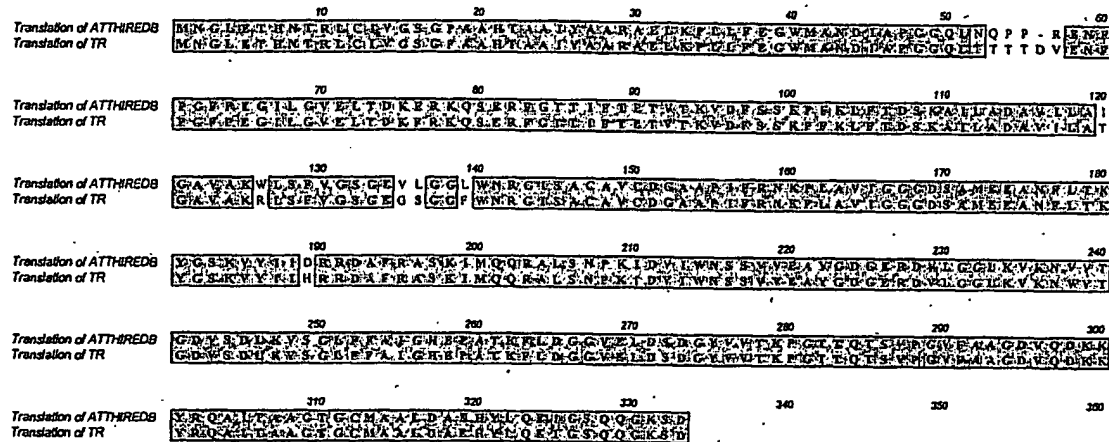


FIGURE 2

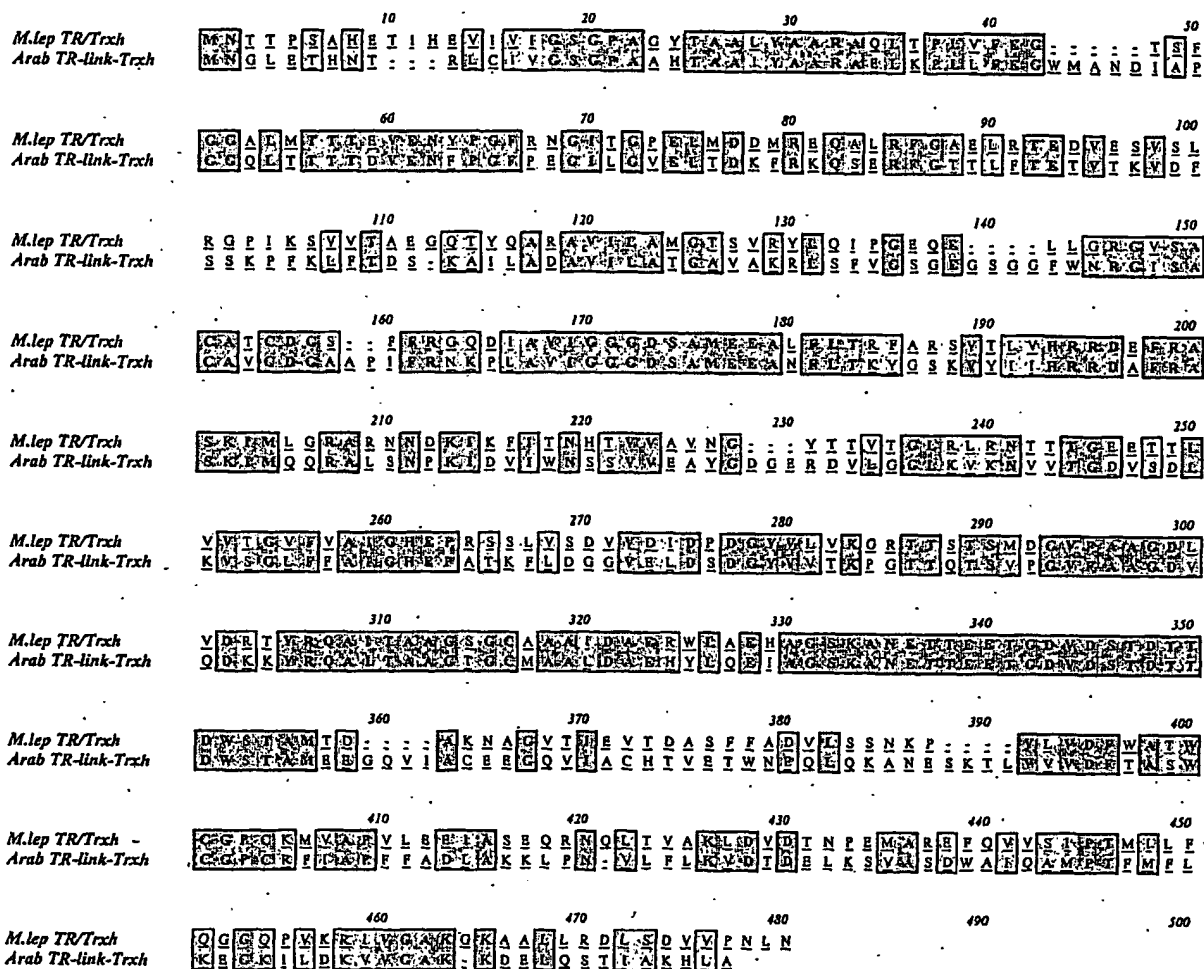
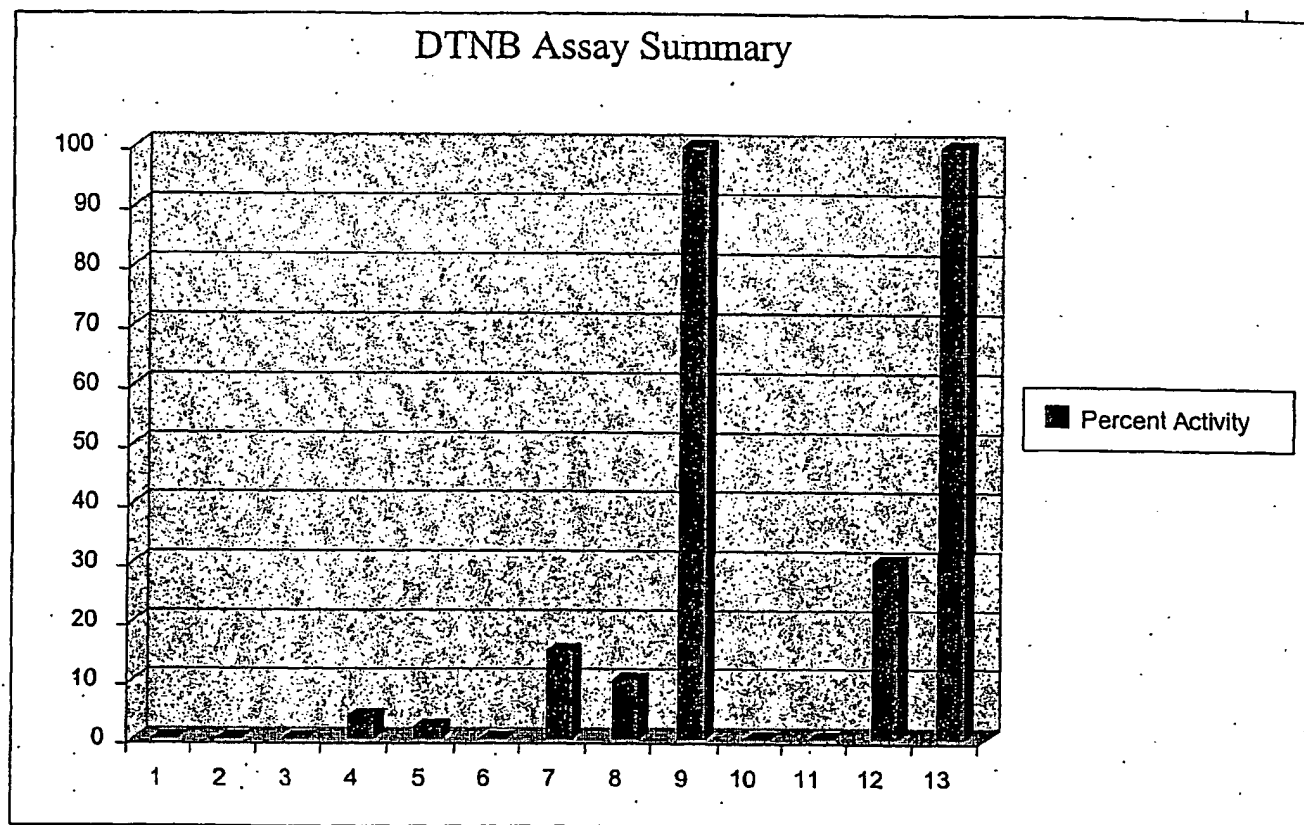


FIGURE 3

**FIGURE 4**

HETEROMULTIMERS

Class	Heteromultimer	Example sequence reference for heteromultimeric subunits
Biosynthetic	3-methyl-2-oxobutanoate dehydrogenase (2-oxoisovalerate dehydrogenase (lipoamide))- E1 component	McKean, <i>et al.</i> Biochim. Biophys. Acta (1992) 1171:109-112 / Chuang, J.L., <i>et al.</i> FEBS Lett. a (1990) 262 (2), 305-309.
Biosynthetic	3-oxoadipate CoA-transferase	Parales, R.E. and Harwood, S.C. J. Bacteriol. (1992) 174:4657-4666
Biosynthetic	anthranilate synthase:indole-3-glycerol phosphate synthase	Zalkin, H.; <i>et al.</i> J. Biol. Chem. (1984) 259:3985-3992.
Biosynthetic	beta-ketoacyl-[acyl carrier protein] synthase I	Siggaard-Andersen, M. <i>et al.</i> Proc. Natl. Acad. Sci. U.S.A. (1991) 88:4114-4118
Biosynthetic	butyrate--acetoacetate CoA-transferase	Fischer, R.J., <i>et al.</i> J. Bacteriol. (1993) 175 (21), 6959-6969.
Biosynthetic	cAMP dependent protein kinase	Mutzel, R. <i>et al.</i> Proc. Natl. Acad. Sci. U.S.A. (1987) 84:6-10 / Burki, E., <i>et al.</i> Gene (1991) 102 (1), 57-65.
Biosynthetic	carbamoyl-phosphate synthase	Shigenobu, S., <i>et al.</i> Nature. (2000) 407 (6800), 81-86.
Biosynthetic	Creatine kinase	Billadello, J.J.; <i>et al.</i> Biochem. Biophys. Res. Commun. (1986) 138:392-398. / Roman, D.; <i>et al.</i> Proc. Natl. Acad. Sci. U.S.A. (1985) 82:8394-8398.
Biosynthetic	gamma-glutamyltransferase (gamma-glutamyl transpeptidase)	Papandrikopoulou, A.; <i>et al.</i> Eur. J. Biochem. (1989) 183:693-698.
Biosynthetic	glutathione transferase	Morrow, C.S. <i>et al.</i> Gene (1989) 75:3-11
Biosynthetic	glycerol-3-phosphate dehydrogenase	Cole, S.T. <i>et al.</i> J. Bacteriol. (1988) 170:2448-2456.
Biosynthetic	guanylate cyclase	Hinsch, K.D. <i>et al.</i> FEBS Lett. (1988) 239:29-34 / Koesling, D. <i>et al.</i> FEBS Lett. (1990) 266:128-132.
Biosynthetic	heterodisulfide reductase	Smith, D.R., <i>et al.</i> J. Bacteriol. (1997) 179 (22), 7135-7155.
Biosynthetic	human cathepsin	Ritonja, A. <i>et al.</i> FEBS Lett. (1988) 228:341-345.
Biosynthetic	Hydrogenase	Menon, N.K. <i>et al.</i> J. Bacteriol. (1990) 172:1969-1977.
Biosynthetic	Meprin A	Johnson, G.D. and Hersh, L.B. J. Biol. Chem. (1992) 267:13505-13512.
Biosynthetic	methionine adenosyltransferase	Horikawa, S.; Tsukada, K. FEBS Lett. (1992) 312:37-41.
Biosynthetic	methylmalonyl-CoA mutase	Jackson, C.A. <i>et al.</i> Gene (1995) 167:127-132.
Biosynthetic	mitochondrial processing peptidase	Pollock, R.A. <i>et al.</i> EMBO J. (1988) 7:3493-3500.
Biosynthetic	Na ⁺ /K ⁺ -exchanging ATPase	Shull, G.E., <i>et al.</i> Biochemistry (1986) 25 (25), 8125-8132. / Mercer, R.W., <i>et al.</i> Mol. Cell. Biol. (1986) 6 (11), 3884-3890 / Mercer, R.W., <i>et al.</i> J. Cell Biol. (1993) 121 (3), 579-586.
Biosynthetic	NAD(+)--dependent isocitrate dehydrogenase	Cupp, J.R. and McAlister-Henn, L. J. Biol. Chem. (1992) 267:16417-16423. / Cupp, J.R. and McAlister-Henn, L. J. Biol. Chem. (1991) 266:22199-22205.
Biosynthetic	phosphoribosylformylglycinamidine synthase	Ebbole, D.J.; Zalkin, H. J. Biol. Chem. (1987) 262:8274-8287.
Biosynthetic	protocatechuate 3,4-dioxygenase	Frazee, R.W.; <i>et al.</i> J. Bacteriol. (1993) 175:6194-6202.
Biosynthetic	S-100 protein	Engelkamp, D.; <i>et al.</i> Biochemistry (1992)

FIGURE 5

		31:10258-10264. / Allore, R.J.; <i>et al.</i> J. Biol. Chem. (1990) 265:15537-15543.
Biosynthetic	sucrose--fructan 6-fructosyltransferase	Sprenger, N.; <i>et al.</i> Proc. Natl. Acad. Sci. U.S.A. (1995) 92:11652-11656.
Biosynthetic	Superoxide dismutase	Capo, C.R.; <i>et al.</i> Biochem. Biophys. Res. Commun. (1990) 173:1186-1193.
Biosynthetic	Urease	Labigne, A.; <i>et al.</i> J. Bacteriol. (1991) 173:1920-1931.
Biosynthetic	urokinase-type plasminogen activator (urokinase)	Belin, D. <i>et al.</i> Eur. J. Biochem. (1985) 148:225-232.
Biosynthetic	methylmalonyl-coenzyme A mutase	Birch, A., <i>et al.</i> J. Bacteriol. (1993) 175 (11), 3511-3519.
Calcium binding	Calcineurin	Muramatsu, T. and Kincaid, R.L. Biochim. Biophys. Acta (1993) 1178 (1), 117-120 / Guerini, D. <i>et al.</i> DNA (1989) 8:675-682.
Calcium binding	Calgranulin	Imamichi, T. <i>et al.</i> Biochem. Biophys. Res. Commun. (1993) 194:819-825.
Calcium binding	Calpain	Aoki, K. <i>et al.</i> FEBS Lett. (1986) 205:313-317.
DNA binding	AP1	van Straaten, F., <i>et al.</i> Proceedings of the National Academy of Sciences of the United States of America. (1983) 80 (11), 3183-3187. / Hattori, K., <i>et al.</i> Proceedings of the National Academy of Sciences of the United States of America. (1988) 85 (23), 9148-9152.
DNA binding	cMyc-Max	Schreiber-Agus, N. <i>et al.</i> Mol. Cell. Biol. (1993) 13 (5), 2765-2775.
DNA binding	DNA binding protein HU-1/HU-2	Laine, B. <i>et al.</i> Eur. J. Biochem. (1980) 103:447-461.
DNA binding	hepatic nuclear factor 1	Bach, I. <i>et al.</i> Nucleic Acids Res. (1992) 20 (16), 4199-4204. / Rey-Campos, J. <i>et al.</i> EMBO J. (1991) 10 (6), 1445-1457.
DNA binding	Integration host factor	Miller, H.I. Cold Spring Harbor symposia on quantitative biology. (1984) 49, 691-698. / Flamm, E. and Weisberg, R.A. J. Mol. Biol. (1985) 183:117-128.
DNA binding	Ku	Reeves, W.H. and Sthoeger, Z.M. J. Biol. Chem. (1989) 264 (9), 5047-5052. / J. Biol. Chem. (1989) 264 (23), 13407-13411.
DNA binding	MutS	Bocker <i>et al.</i> 1999. Cancer Research 59, 816-822.
DNA binding	NF-E2	Chan, J.Y. <i>et al.</i> Proc. Natl. Acad. Sci. U.S.A. (1993) 90 (23), 11366-11370. / Toki, T., <i>et al.</i> Oncogene (1997) 14 (16), 1901-1910.
DNA binding	nuclear factor kB (NFkB)	Kieran M, <i>et al.</i> Cell. (1990) Sep 7;62(5):1007-18. / Ruben SM, <i>et al.</i> Science (1991) Mar 22;251(5000):1490-3. Erratum in: Science (1991) Oct 4;254(5028):11
Electron transport	corrinoid/iron-sulfur protein	Lu, W.P. <i>et al.</i> J. Biol. Chem. (1993) 268:5605-5614.
Electron transport	cytochrome d ubiquinol oxidase	Green, G.N. <i>et al.</i> J. Biol. Chem. (1988) 263:13138-13143.
Electron transport	cytochrome-c3 hydrogenase	Menon, N.K. <i>et al.</i> J. Bacteriol. (1987) 169:5401-5407.
Electron transport	electron transfer flavoprotein	Finocchiaro, G. <i>et al.</i> Biol. Chem. (1988) 263:15773-15780. / Finocchiaro, G. <i>et al.</i> Eur. J. Biochem. (1993) 213:1003-1008.

Electron transport	xylene monooxygenase	Shaw, J.P. and Harayama, S. <i>Eur. J. Biochem.</i> (1992) 209:51-61. / Kasai, Y., <i>et al.</i> <i>J. Bacteriol.</i> (2001) 183 (22), 6662-6666.
Growth factor	hepatocyte growth factor	Nakamura, T. <i>et al.</i> <i>Nature</i> (1989) 342:440-443.
Growth factor	human chorionic gonadotropin	Morgan, F.J. <i>et al.</i> <i>J. Biol. Chem.</i> (1975) 250 (13), 5247-5258.
Growth factor	Platelet-derived growth factor	Takimoto, Y., <i>et al.</i> <i>Hiroshima J. Med. Sci.</i> (1993) 42 (1), 47-52. / Josephs, S.F., <i>et al.</i> <i>Science</i> (1984) 225 (4662), 636-639.
Hormone	Bombyxin	Adachi, T. <i>et al.</i> <i>J. Biol. Chem.</i> (1989) 264:7681-7685.
Hormone	Follicle stimulating hormone	Fiddes, J.C. and Goodman, H.M. <i>J. Mol. Appl. Genet.</i> (1981) 1 (1), 3-18. / Watkins, P.C., <i>et al.</i> <i>DNA</i> (1987) 6 (3), 205-212.
Hormone	Insulin	Bell, G.I., Pictet, R.L., Rutter, W.J., Cordell, B., Tischer, E. and Goodman, H.M. Sequence of the human insulin gene. <i>Nature</i> . 284 (5751), 26-32 (1980)
Hormone	Luteinizing Hormone	Fiddes, J.C. and Goodman, H.M. <i>J. Mol. Appl. Genet.</i> (1981) 1 (1), 3-18. / Shome, B. and Parlow, A.F. <i>J. Clin. Endocrinol. Metab.</i> (1973) 36 (3), 618-621.
Hormone	Thyroid stimulating hormone	Fiddes, J.C. and Goodman, H.M. <i>J. Mol. Appl. Genet.</i> (1981) 1 (1), 3-18. / Hayashizaki Y, <i>et al.</i> <i>FEBS Lett.</i> (1985) 188 (2), 394-400.
Immune	B-cell antigen receptor complex	Hashimoto, S. <i>et al.</i> <i>J. Immunol.</i> (1993) 150 (2), 491-498. / Flaswinkel, H. and Reth, M. <i>Immunogenetics</i> (1992) 36 (4), 266-269.
Immune	Cell surface CD8 molecules	Ureta-Vidal, A., <i>et al.</i> <i>Immunogenetics</i> (1999) 49 (7-8), 718-721.
Immune	human complement subcomponent C1q	Sellar, G.C. <i>et al.</i> <i>Biochem. J.</i> (1991) 274:481-490.
Immune	T cell receptor	Talke, B.L. <i>et al.</i> <i>Scand. J. Immunol.</i> (2001) 54 (1-2), 204-210.
Photosynthesis	C-phycocyanin	Offner, G.D. <i>et al.</i> <i>J. Biol. Chem.</i> (1981) 256:12167-12175. / Troxler, R.F. <i>et al.</i> <i>J. Biol. Chem.</i> (1981) 256:12176-12184.
Photosynthesis	ferredoxin-thioredoxin reductase	Chow, L.P. <i>et al.</i> <i>Eur. J. Biochem.</i> (1995) 231:149-156. / Iwadate, H. <i>et al.</i> <i>Eur. J. Biochem.</i> (1994) 223:465-471.
Photosynthesis	Light harvesting complex I	<i>Proc. Natl. Acad. Sci. U.S.A.</i> (1984) 81, 189-192.
Photosynthetic	cytochrome b559	Carrillo, N. <i>et al.</i> <i>Curr Genet.</i> 1986;10(8):619-24.
Protease	ATP-dependent Clp protease	Gerth, U. <i>et al.</i> <i>Gene</i> (1996) 181:77-83. / Kunst, F. <i>et al.</i> <i>Nature</i> (1997) 390 (6657), 249-256.
Receptor	alpha-2-macroglobulin receptor	Strickland, D.K. <i>et al.</i> <i>J. Biol. Chem.</i> (1990) 265:17401-17404. / Strickland, D.K. <i>et al.</i> <i>J. Biol. Chem.</i> (1991) 266:13364-13369.
Receptor	Interleukin-2 receptor	Ishida, N. <i>et al.</i> <i>Nucleic Acids Res.</i> (1985) 13:7579-7589. / Hatakeyama, M. <i>et al.</i> <i>Science</i> (1989) 244:551-556 / Takeshita, T. <i>et al.</i> <i>Science</i> (1992) 257:379-382.
Receptor	platelet-derived growth factor receptor	Lee, K.H. <i>et al.</i> <i>Mol. Cell. Biol.</i> (1990) 10:2237-2246. / Herren, B. <i>et al.</i> <i>Biochim. Biophys. Acta</i> 1173 (3), 294-302 (1993).
Structural	Hemoglobin	Heindell, H.C. <i>et al.</i> <i>Cell</i> (1978) 15 (1), 43-54. /

		Best, J.S. <i>et al.</i> Hoppe-Seyler's Z. Physiol. Chem. (1989) 350 (5), 563-580. / Hardison, R.C. J. Biol. Chem. (1981) 256 (22), 11780-11786.
Structural	human platelet glycoprotein Ib	Wenger, R.H. <i>et al.</i> Biochem. Biophys. Res. Commun. (1988) 156 (1), 389-395. / Yagi, M. <i>et al.</i> J. Biol. Chem. (1994) 269 (26), 17424-17427.
Structural	Plasma fibronectin	Kornblihtt, A.R. <i>et al.</i> Proc. Natl. Acad. Sci. U.S.A. (1983) 80:3218-3222.
Structural	Spectrin	Sahr, K.E. <i>et al.</i> J. Biol. Chem. (1990) 265:4434-4443. / Winkelmann, J.C. <i>et al.</i> J. Biol. Chem. (1990) 265:11827-11832.
Structural	Tubulin	Ponstingl, H. <i>et al.</i> Proc. Natl. Acad. Sci. U.S.A. (1981) 78:2757-2761. / Krauhs, E. <i>et al.</i> Proc. Natl. Acad. Sci. U.S.A. (1981) 78:4156-4160.
Toxin	Agkisacutacin	Cheng, X. <i>et al.</i> Biochem. Biophys. Res. Commun. (1999) 265 (2), 530-535.
Toxin	Beta bungarotoxins	Kondo, K. <i>et al.</i> J. Biochem. (1978) 83:101-115.
Toxin	Crotoxin	Bouchier, C. <i>et al.</i> Nucleic Acids Res. (1988) 16 (18), 9050.
Toxin	Mojave toxin	John, T.R. <i>et al.</i> Gene (1994) 139:229-234.
Toxin	venom protein C9S3	Rowan, E.G. <i>et al.</i> Nucleic Acids Res. (1990) 18:1639. / Joubert, F.J. and Viljoen, C.C. Hoppe-Seyler's Z. Physiol. Chem. (1979) 360:1075-1090.
Miscellaneous	Inhibin	Forage, R.G. <i>et al.</i> Proc. Natl. Acad. Sci. U.S.A. (1986) 83:3091-3095.
Miscellaneous	Monellin	Frank, G. and Zuber, H. Hoppe-Seyler's Z. Physiol. Chem. (1976) 357:585-592.
Miscellaneous	mRNA capping enzyme	Niles, E.G. <i>et al.</i> , J. Virology (1986) 153:96-112.
Miscellaneous	Soybean insulin-binding protein si30	Barbashov, S.F. <i>et al.</i> Bioorg. Khim. (1991) 17:421-423.

SEQUENCE LISTING

<110> SemBioSys Genetics, Inc.
Syngenta Participations AG

<120> METHODS FOR THE PRODUCTION OF MULTIMERIC PROTEINS, AND RELATED COMPOSITIONS

<130> 38814-351PC

<140> Not Yet Assigned

<141> Herewith

<160> 313

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 1
taccatggct tcggaagaag ga 22

<210> 2
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 2
gaaagcttaa gccaaagtgt tg 22

<210> 3
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 3
ggccagcaca ctaccatgaa tggctctcgaa actcac 36

<210> 4
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 4
ttaagcttca atcactctta ccttgctg 28

<210> 5
<211> 72
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer

<400> 5

actggagatg ttgactcgac ggatactacg gattgggtcga cggctatgga agaaggacaa 60
gtgatcgct gc 72

<210> 6

<211> 80

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 6

atccgtcgag tcaacatctc cagtttcttc ggtgggtctcg ttagccttcg atccagcaat 60
ctcttgtaag aatgctctgc 80

<210> 7

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 7

gtggaagctt atggagatgg ag 22

<210> 8

<211> 1002

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric

<400> 8

atgaatgggtc	tcgaaactca	caacacaagg	ctctgtatcg	taggaagtgg	cccagcggca	60
cacacggcgg	cgatttacgc	agctagggct	gaacttaaac	ctcttctctt	cgaaggatgg	120
atggctaacc	acatcgctcc	cggtgggtcaa	ctaacaacca	ccaccgacgt	cgagaatttc	180
cccggaatttc	cagaaggtat	tctcggagta	gagctcactg	acaaattccg	taaacaatcg	240
gagcgatttcg	gtactacgat	atttacagag	acgggtgacga	aagtcgattt	ctcttcgaaa	300
ccgtttaagc	tattcacaga	ttcaaaaagcc	attctcgctg	acgctgtgat	tctcgctact	360
ggagctgtgg	ctaagcggct	tagcttcggt	ggatctgggtg	aaggttctgg	aggtttctgg	420
aaccgtggaa	tctccgcttg	tgctgtttgc	gacggagctg	ctccgatatt	ccgtaacaaa	480
cctcttgccg	tgatcgggtg	aggcgattca	gcaatggaag	aagcaaactt	tcttacaaaa	540
tatggatcta	aagtgtatat	aatccatagg	agagatgctt	ttagagcgtc	taagattatg	600
cagcagcgag	ctttgtctaa	tcctaagatt	gatgtgattt	ggaactcgtc	tgttggtgaa	660
gcttatggag	atggagaaa	agatgtgctt	ggaggattga	aagtgaagaa	tgtggttacc	720
ggagatgttt	ctgatttaaa	agtttctgga	ttgttctttg	ctattggtca	tgagccagct	780
accaagtttt	tggatgggtg	tgttgagtta	gattcggatg	gttatgttgt	cacgaagcct	840
ggtactacac	agactagcgt	tcccggagtt	ttcgtgcggg	gtgatgttca	ggataagaag	900
tataggcaag	ccatcactgc	tgcaaggaact	gggtgcatgg	cagctttgga	tgcagagcat	960
tacttacaag	agattggatc	tcagcaagggt	aagagtgatt	ga		1002

<210> 9

<211> 999

<212> DNA

<213> Arabidopsis thaliana

<400> 9

atgaatgggtc	tcgaaactca	caacacaagg	ctctgtatcg	taggaagtgg	cccagcggca	60
cacacggcgg	cgatttacgc	agctagggct	gaacttaaac	ctcttctctt	cgaaggatgg	120
atggctaacc	acatcgctcc	cggtgggtcaa	ctaacaacca	caccgcgtga	gaatttcccc	180
ggatttccag	aaggatttct	cggagtagag	ctcactgaca	aattccgtaa	acaatcggag	240
cgattcggta	ctacgatatt	tacagagacg	gtgacgaaag	tcgatttctc	ttcgaaaccg	300

```

ttaaagctat tcacagattc aaaagccatt ctcgctgacg ctgtgattct cgctatcgga 360
gctgtggcta agtggcttag ctctcggtgga tctgggtgaag ttctcggagg tttgtggaac 420
cgtggaatct ccgcttgtgc tgtttgcgac ggagctgctc cgatattccg caacaaacct 480
cttgcggtga tcggtggagg cgattctgca atggaagaag caaactttct tacaaaatat 540
ggatctaaag tgatatataat cgataggaga gatgctttta gagcgtctaa gattatgcag 600
cagcgagctt tgtctaattc taagattgat gtgatttgga actcgtctgt tgtggaagct 660
tatggagatg gagaaagaga tgtgcttgga ggattgaaag tgaagaatgt ggttaccgga 720
gatgtttctg atttaaaagt ttctggattg ttctttgcta ttggtcatga gccagctacc 780
aagtttttgg atggtgggtg tgagttagat tcggatgggt atgttggtcac gaagcctggg 840
actacacaga ctacggttcc cggagttttc gctgcgggtg atgttcagga taagaagtat 900
aggcaagcca tcactgctgc aggaactggg tgcattggcg ctttggatgc agagcattac 960
ttacaagaga ttgatctca gcaaggtaag agtgattga 999

```

<210> 10

<211> 1002

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric

<221> CDS

<222> (1)...(1002)

<223> cDNA encoding NADPH thioredoxin reductase

<400> 10

```

atg aat ggt ctc gaa act cac aac aca agg ctc tgt atc gta gga agt 48
Met Asn Gly Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser
1 5 10 15

```

```

ggc cca gcg gca cac acg gcg gcg att tac gca gct agg gct gaa ctt 96
Gly Pro Ala Ala His Thr Ala Ala Ile Tyr Ala Ala Arg Ala Glu Leu
20 25 30

```

```

aaa cct ctt ctc ttc gaa gga tgg atg gct aac gac atc gct ccc ggt 144
Lys Pro Leu Leu Phe Glu Gly Trp Met Ala Asn Asp Ile Ala Pro Gly
35 40 45

```

```

ggg caa cta aca acc acc acc gac gtc gag aat ttc ccc gga ttt cca 192
Gly Gln Leu Thr Thr Thr Asp Val Glu Asn Phe Pro Gly Phe Pro
50 55 60

```

```

gaa ggt att ctc gga gta gag ctc act gac aaa ttc cgt aaa caa tcg 240
Glu Gly Ile Leu Gly Val Glu Leu Thr Asp Lys Phe Arg Lys Gln Ser
65 70 75 80

```

```

gag cga ttc ggt act acg ata ttt aca gag acg gtg acg aaa gtc gat 288
Glu Arg Phe Gly Thr Thr Ile Phe Thr Glu Thr Val Thr Lys Val Asp
85 90 95

```

```

ttc tct tcg aaa ccg ttt aag cta ttc aca gat tca aaa gcc att ctc 336
Phe Ser Ser Lys Pro Phe Lys Leu Phe Thr Asp Ser Lys Ala Ile Leu
100 105 110

```

```

gct gac gct gtg att ctc gct act gga gct gtg gct aag cgg ctt agc 384
Ala Asp Ala Val Ile Leu Ala Thr Gly Ala Val Ala Lys Arg Leu Ser
115 120 125

```

```

ttc gtt gga tct ggt gaa ggt tct gga ggt ttc tgg aac cgt gga atc 432
Phe Val Gly Ser Gly Glu Gly Ser Gly Gly Phe Trp Asn Arg Gly Ile
130 135 140

```

```

tcc gct tgt gct gtt tgc gac gga gct gct ccg ata ttc cgt aac aaa 480
Ser Ala Cys Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg Asn Lys
145 150 155 160

```

```

cct ctt gcg gtg atc ggt gga ggc gat tca gca atg gaa gaa gca aac 528

```

Pro	Leu	Ala	Val	Ile	Gly	Gly	Gly	Asp	Ser	Ala	Met	Glu	Glu	Ala	Asn		
				165					170						175		
ttt	ctt	aca	aaa	tat	gga	tct	aaa	gtg	tat	ata	atc	cat	agg	aga	gat	576	
Phe	Leu	Thr	Lys	Tyr	Gly	Ser	Lys	Val	Tyr	Ile	Ile	His	Arg	Arg	Asp		
			180					185					190				
gct	ttt	aga	gcg	tct	aag	att	atg	cag	cag	cga	gct	ttg	tct	aat	cct	624	
Ala	Phe	Arg	Ala	Ser	Lys	Ile	Met	Gln	Gln	Arg	Ala	Leu	Ser	Asn	Pro		
		195					200					205					
aag	att	gat	gtg	att	tgg	aac	tcg	tct	gtt	gtg	gaa	gct	tat	gga	gat	672	
Lys	Ile	Asp	Val	Ile	Trp	Asn	Ser	Ser	Val	Val	Glu	Ala	Tyr	Gly	Asp		
	210					215					220						
gga	gaa	aga	gat	gtg	ctt	gga	gga	ttg	aaa	gtg	aag	aat	gtg	gtt	acc	720	
Gly	Glu	Arg	Asp	Val	Leu	Gly	Gly	Leu	Lys	Val	Lys	Asn	Val	Val	Thr		
225					230					235					240		
gga	gat	gtt	tct	gat	tta	aaa	gtt	tct	gga	ttg	ttc	ttt	gct	att	ggg	768	
Gly	Asp	Val	Ser	Asp	Leu	Lys	Val	Ser	Gly	Leu	Phe	Phe	Ala	Ile	Gly		
				245					250					255			
cat	gag	cca	gct	acc	aag	ttt	ttg	gat	ggg	ggg	gtt	gag	tta	gat	tcg	816	
His	Glu	Pro	Ala	Thr	Lys	Phe	Leu	Asp	Gly	Gly	Val	Glu	Leu	Asp	Ser		
			260					265					270				
gat	ggg	tat	gtt	gtc	acg	aag	cct	ggg	act	aca	cag	act	agc	gtt	ccc	864	
Asp	Gly	Tyr	Val	Val	Thr	Lys	Pro	Gly	Thr	Thr	Gln	Thr	Ser	Val	Pro		
		275					280					285					
gga	gtt	ttc	gct	gcg	ggg	gat	gtt	cag	gat	aag	aag	tat	agg	caa	gcc	912	
Gly	Val	Phe	Ala	Ala	Gly	Asp	Val	Gln	Asp	Lys	Lys	Tyr	Arg	Gln	Ala		
	290					295					300						
atc	act	gct	gca	gga	act	ggg	tgc	atg	gca	gct	ttg	gat	gca	gag	cat	960	
Ile	Thr	Ala	Ala	Gly	Thr	Gly	Cys	Met	Ala	Ala	Leu	Asp	Ala	Glu	His		
305					310					315					320		
tac	tta	caa	gag	att	gga	tct	cag	caa	ggg	aag	agt	gat	tga			1002	
Tyr	Leu	Gln	Glu	Ile	Gly	Ser	Gln	Gln	Gly	Lys	Ser	Asp	*				
				325					330								

<210> 11

<211> 333

<212> PRT

<213> Artificial Sequence

<220>

<223> Chimeric

<400> 11

Met	Asn	Gly	Leu	Glu	Thr	His	Asn	Thr	Arg	Leu	Cys	Ile	Val	Gly	Ser		
1				5					10					15			
Gly	Pro	Ala	Ala	His	Thr	Ala	Ala	Ile	Tyr	Ala	Ala	Arg	Ala	Glu	Leu		
		20						25					30				
Lys	Pro	Leu	Leu	Phe	Glu	Gly	Trp	Met	Ala	Asn	Asp	Ile	Ala	Pro	Gly		
		35					40					45					
Gly	Gln	Leu	Thr	Thr	Thr	Thr	Asp	Val	Glu	Asn	Phe	Pro	Gly	Phe	Pro		
	50					55					60						
Glu	Gly	Ile	Leu	Gly	Val	Glu	Leu	Thr	Asp	Lys	Phe	Arg	Lys	Gln	Ser		
65					70					75					80		
Glu	Arg	Phe	Gly	Thr	Thr	Ile	Phe	Thr	Glu	Thr	Val	Thr	Lys	Val	Asp		
				85					90					95			
Phe	Ser	Ser	Lys	Pro	Phe	Lys	Leu	Phe	Thr	Asp	Ser	Lys	Ala	Ile	Leu		

```

      100      105      110
Ala Asp Ala Val Ile Leu Ala Thr Gly Ala Val Ala Lys Arg Leu Ser
      115      120      125
Phe Val Gly Ser Gly Glu Gly Ser Gly Gly Phe Trp Asn Arg Gly Ile
      130      135      140
Ser Ala Cys Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg Asn Lys
145      150      155      160
Pro Leu Ala Val Ile Gly Gly Gly Asp Ser Ala Met Glu Glu Ala Asn
      165      170      175
Phe Leu Thr Lys Tyr Gly Ser Lys Val Tyr Ile Ile His Arg Arg Asp
      180      185      190
Ala Phe Arg Ala Ser Lys Ile Met Gln Gln Arg Ala Leu Ser Asn Pro
      195      200      205
Lys Ile Asp Val Ile Trp Asn Ser Ser Val Val Glu Ala Tyr Gly Asp
      210      215      220
Gly Glu Arg Asp Val Leu Gly Gly Leu Lys Val Lys Asn Val Val Thr
225      230      235      240
Gly Asp Val Ser Asp Leu Lys Val Ser Gly Leu Phe Phe Ala Ile Gly
      245      250      255
His Glu Pro Ala Thr Lys Phe Leu Asp Gly Gly Val Glu Leu Asp Ser
      260      265      270
Asp Gly Tyr Val Val Thr Lys Pro Gly Thr Thr Gln Thr Ser Val Pro
      275      280      285
Gly Val Phe Ala Ala Gly Asp Val Gln Asp Lys Lys Tyr Arg Gln Ala
      290      295      300
Ile Thr Ala Ala Gly Thr Gly Cys Met Ala Ala Leu Asp Ala Glu His
305      310      315      320
Tyr Leu Gln Glu Ile Gly Ser Gln Gln Gly Lys Ser Asp
      325      330

```

<210> 12

<211> 332

<212> PRT

<213> Arabidopsis thaliana

<400> 12

```

Met Asn Gly Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser
  1      5      10      15
Gly Pro Ala Ala His Thr Ala Ala Ile Tyr Ala Ala Arg Ala Glu Leu
      20      25      30
Lys Pro Leu Leu Phe Glu Gly Trp Met Ala Asn Asp Ile Ala Pro Gly
      35      40      45
Gly Gln Leu Asn Gln Pro Pro Arg Glu Asn Phe Pro Gly Phe Pro Glu
      50      55      60
Gly Ile Leu Gly Val Glu Leu Thr Asp Lys Phe Arg Lys Gln Ser Glu
65      70      75      80
Arg Phe Gly Thr Thr Ile Phe Thr Glu Thr Val Thr Lys Val Asp Phe
      85      90      95
Ser Ser Lys Pro Phe Lys Leu Phe Thr Asp Ser Lys Ala Ile Leu Ala
      100      105      110
Asp Ala Val Ile Leu Ala Ile Gly Ala Val Ala Lys Trp Leu Ser Phe
      115      120      125
Val Gly Ser Gly Glu Val Leu Gly Gly Leu Trp Asn Arg Gly Ile Ser
      130      135      140
Ala Cys Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg Asn Lys Pro
145      150      155      160
Leu Ala Val Ile Gly Gly Gly Asp Ser Ala Met Glu Glu Ala Asn Phe
      165      170      175
Leu Thr Lys Tyr Gly Ser Lys Val Tyr Ile Ile Asp Arg Arg Asp Ala
      180      185      190
Phe Arg Ala Ser Lys Ile Met Gln Arg Ala Leu Ser Asn Pro Lys
      195      200      205
Ile Asp Val Ile Trp Asn Ser Ser Val Val Glu Ala Tyr Gly Asp Gly
      210      215      220
Glu Arg Asp Val Leu Gly Gly Leu Lys Val Lys Asn Val Val Thr Gly
225      230      235      240

```



```

Asp Val Ser Asp Leu Lys Val Ser Gly Leu Phe Phe Ala Ile Gly His
      245      250      255
Glu Pro Ala Thr Lys Phe Leu Asp Gly Gly Val Glu Leu Asp Ser Asp
      260      265      270
Gly Tyr Val Val Thr Lys Pro Gly Thr Thr Gln Thr Ser Val Pro Gly
      275      280      285
Val Phe Ala Ala Gly Asp Val Gln Asp Lys Lys Tyr Arg Gln Ala Ile
      290      295      300
Thr Ala Ala Gly Thr Gly Cys Met Ala Ala Leu Asp Ala Glu His Tyr
305      310      315      320
Leu Gln Glu Ile Gly Ser Gln Gln Gly Lys Ser Asp
      325      330

```

<210> 13
 <211> 333
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Chimeric

```

<400> 13
Met Asn Gly Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser
 1      5      10      15
Gly Pro Ala Ala His Thr Ala Ala Ile Tyr Ala Ala Arg Ala Glu Leu
      20      25      30
Lys Pro Leu Leu Phe Glu Gly Trp Met Ala Asn Asp Ile Ala Pro Gly
      35      40      45
Gly Gln Leu Thr Thr Thr Thr Asp Val Glu Asn Phe Pro Gly Phe Pro
      50      55      60
Glu Gly Ile Leu Gly Val Glu Leu Thr Asp Lys Phe Arg Lys Gln Ser
65      70      75      80
Glu Arg Phe Gly Thr Thr Ile Phe Thr Glu Thr Val Thr Lys Val Asp
      85      90      95
Phe Ser Ser Lys Pro Phe Lys Leu Phe Thr Asp Ser Lys Ala Ile Leu
      100      105      110
Ala Asp Ala Val Ile Leu Ala Thr Gly Ala Val Ala Lys Arg Leu Ser
      115      120      125
Phe Val Gly Ser Gly Glu Gly Ser Gly Gly Phe Trp Asn Arg Gly Ile
      130      135      140
Ser Ala Cys Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg Asn Lys
145      150      155      160
Pro Leu Ala Val Ile Gly Gly Gly Asp Ser Ala Met Glu Glu Ala Asn
      165      170      175
Phe Leu Thr Lys Tyr Gly Ser Lys Val Tyr Ile Ile His Arg Arg Asp
      180      185      190
Ala Phe Arg Ala Ser Lys Ile Met Gln Gln Arg Ala Leu Ser Asn Pro
      195      200      205
Lys Ile Asp Val Ile Trp Asn Ser Ser Val Val Glu Ala Tyr Gly Asp
      210      215      220
Gly Glu Arg Asp Val Leu Gly Gly Leu Lys Val Lys Asn Val Val Thr
225      230      235      240
Gly Asp Val Ser Asp Leu Lys Val Ser Gly Leu Phe Phe Ala Ile Gly
      245      250      255
His Glu Pro Ala Thr Lys Phe Leu Asp Gly Gly Val Glu Leu Asp Ser
      260      265      270
Asp Gly Tyr Val Val Thr Lys Pro Gly Thr Thr Gln Thr Ser Val Pro
      275      280      285
Gly Val Phe Ala Ala Gly Asp Val Gln Asp Lys Lys Tyr Arg Gln Ala
      290      295      300
Ile Thr Ala Ala Gly Thr Gly Cys Met Ala Ala Leu Asp Ala Glu His
305      310      315      320
Tyr Leu Gln Glu Ile Gly Ser Gln Gln Gly Lys Ser Asp
      325      330

```

<210> 14
 <211> 3129
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> CDS
 <222> (1555) ... (1899)

<223> Chimeric

<400> 14
 ctgcaggaat tcattgtact cccagtatca ttatagttaa agttttggct ctctcgccgg 60
 tggtttttta cctctatttta aagggggttt ccacctaaaa attctgggtat cattctcact 120
 ttactttgta ctttaatttc tcataatcct tgggttgaat taccacgctt ccgcacacga 180
 tatccctaca aattttattat ttgttaaaca ttttcaaacc gcataaaatt ttatgaagtc 240
 ccgtctatct ttaattgtagt ctaacatttt catattgaaa tatataattt acttaatttt 300
 agcgttggta gaaagcataa tgatttatcc ttattcttct tcatataaat gtttaataata 360
 caatataaac aaattcttta ccttaagaag gatttcccat tttatatttt aaaaatataat 420
 ttatcaaata tttttcaacc acgtaaatct cataataata agttgtttca aaagtaataa 480
 aatttaactc cataattttt ttattcgact gatcttaaag caacaccag tgacacaact 540
 agccattttt ttctttgaat aaaaaaatcc aattatcatt gtattttttt tatacaatga 600
 aaatttcacc aaacaatcat ttgtgggtatt tctgaagcaa gtcattgttat gcaaaattct 660
 ataattccca tttagacta cggaagtaac tgaagatctg cttttacatg cgagacacat 720
 cttctaaagt aattttaata atagttacta tattcaagat ttcataatc aaataactcaa 780
 tattacttct aaaaaattaa ttagatataa ttaaaatatt acttttttaa ttttaagttt 840
 aattgttgaa ttgttgacta ttgatttatt attctactat gtttaaatg ttttatagat 900
 agtttaaagt aaatataagt aatgtagtag agtggttagag tgttacccta aaccataaac 960
 tataagattt atgggtggact aattttcata tatttcttat tgcttttacc ttttcttggg 1020
 atgtaagtcc gtaactggaa ttactgtggg ttgccatggc actctgtggg cttttgggtt 1080
 atgcatggat gcttgcgcaa gaaaaagaca aagaacaaag aaaaaagaca aaacagagag 1140
 acaaaacgca atcacacaac caactcaaat tagtcaactg ctgatcaaga tcgccgcgtc 1200
 catgtatgtc taaatgccat gcaaagcaac acgtgcttaa catgcacttt aaatggctca 1260
 cccatctcaa cccacacaca aacacattgc ctttttcttc atcatcacca caaccacctg 1320
 tataatattc ttctcttccg ccacctcaat ttcttcaact caacacacgt caacctgcat 1380
 atgcgtgtca tcccatgccc aaatctccat gcatgttcca accaccttct ctcttatata 1440
 atacctataa atacctctaa tatcactcac ttcttctcat atccatccat ccagagtact 1500
 actactctac tactataata cccaaccca actcatattc aatactactc tact atg 1557
 Met
 1
 gct tgc gaa gaa gga caa gtg atc gcc tgc cac acc gtt gag aca tgg 1605
 Ala Ser Glu Glu Gly Gln Val Ile Ala Cys His Thr Val Glu Thr Trp
 5 10 15
 aac gag cag ctt cag aag gct aat gaa tcc aaa act ctt gtg gtg gtt 1653
 Asn Glu Gln Leu Gln Lys Ala Asn Glu Ser Lys Thr Leu Val Val Val
 20 25 30
 gat ttc acg gct tct tgg tgt gga cca tgt cgt ttc atc gct cca ttc 1701
 Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro Phe
 35 40 45
 ttt gct gat ttg gct aag aaa ctt cct aac gtg ctt ttc ctc aag gtt 1749
 Phe Ala Asp Leu Ala Lys Lys Leu Pro Asn Val Leu Phe Leu Lys Val
 50 55 60 65
 gat act gat gaa ttg aag tgc gtg gca agt gat tgg gcg ata cag gcg 1797
 Asp Thr Asp Glu Leu Lys Ser Val Ala Ser Asp Trp Ala Ile Gln Ala
 70 75 80
 atg cca acc ttc atg ttt ttg aag gaa ggg aag att ttg gac aaa gtt 1845
 Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Leu Asp Lys Val
 85 90 95
 gtt gga gcc aag aaa gat gag ctt cag tct acc att gcc aaa cac ttg 1893
 Val Gly Ala Lys Lys Asp Glu Leu Gln Ser Thr Ile Ala Lys His Leu

100

105

110

gct taa gcttaataag tatgaactaa aatgcatgta ggtgtaagag ctcatggaga 1949
Ala *

```

gcatggaata ttgtatccga ccatgtaaca gtataataac tgagctccat ctccacttctt 2009
ctatgaataa acaaaggatg ttatgatata ttaacactct atctatgcac cttattgttc 2069
tatgataaat ttcctcttat tattataaat catctgaatc gtgacggctt atggaatgct 2129
tcaaatagta caaaaacaaa tgtgtactat aagactttct aaacaattct aacttttagca 2189
ttgtgaacga gacataagtg ttaagaagac ataacaatta taatggaaga agtttgtctc 2249
catttatata ttatatatta cccacttatg tattatatta ggatgttaag gagacataac 2309
aattataaag agagaagttt gtatccattt atatattata tactacccat ttatatatta 2369
tacttatcca cttattttaat gtctttataa ggtttgatcc atgatatttc taatatttta 2429
gttgatatgt atatgaaagg gtactatttg aactctctta ctctgtataa aggttggatc 2489
atccttaaag tgggtctatt taattttatt gcttcttaca gataaaaaaa aaattatgag 2549
ttggtttgat aaaaatttga aggatttaaa ataataataa ataataaata acatataata 2609
tatgtatata aattttattat aatataacat ttatctataa aaaagtaaatt attgtcataa 2669
atctatacaa tcgttttagcc ttgctggacg actctcaatt atttaaacga gagtaaaccat 2729
atgtgacttt ttggttattt aacaaattat tatttaaacac tatatgaaat tttttttttt 2789
tatcggaag gaaataaaat taaattagga gggacaatgg tgtgtcccaa tccttatata 2849
accaacttcc acaggaaggt caggtcgggg acaacaaaaa aacaggcaag ggaaattttt 2909
taatttgggt tgtcttgttt gctgcataat ttatgcagta aaacactaca cataaccctt 2969
ttagcagtag agcaatggtt gaccgtgtgc ttagcttctt ttattttatt tttttatcag 3029
caaagaataa ataaaataaa atgagacact tcagggatgt ttcaaccctt atacaaaaacc 3089
ccaaaaacaa gtttcctagc accctaccaa ctaagggtacc 3129

```

<210> 15

<211> 114

<212> PRT

<213> Artificial Sequence

<220>

<223> Chimeric

<400> 15

```

Met Ala Ser Glu Glu Gly Gln Val Ile Ala Cys His Thr Val Glu Thr
 1           5           10          15
Trp Asn Glu Gln Leu Gln Lys Ala Asn Glu Ser Lys Thr Leu Val Val
          20          25          30
Val Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro
          35          40          45
Phe Phe Ala Asp Leu Ala Lys Lys Leu Pro Asn Val Leu Phe Leu Lys
          50          55          60
Val Asp Thr Asp Glu Leu Lys Ser Val Ala Ser Asp Trp Ala Ile Gln
          65          70          75          80
Ala Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Leu Asp Lys
          85          90          95
Val Val Gly Ala Lys Lys Asp Glu Leu Gln Ser Thr Ile Ala Lys His
          100         105         110
Leu Ala

```

<210> 16

<211> 3888

<212> DNA

<213> Artificial sequence

<220>

<223> Chimeric

<221> CDS

<222> (1555) ... (1907)

<221> CDS

<222> (2148) ... (2659)

<400> 16

```

ctgcaggaat tcattgtact cccagtatca ttatagttaa agttttggct ctctcgccgg 60
tggtttttta cctctattta aaggggtttt ccacctaata attctggtat cattctcact 120
ttacttgtta ctttaatttc tcataatcct tggttgaaat tatcacgctt ccgcacacga 180
tatccctaca aattttattat ttgttaaaca ttttcaaacc gcataaaatt ttatgaagtc 240
ccgtctatct ttaatgtagt ctaacatttt catattgaaa tatataattt acttaatttt 300
agcgttggta gaaagcataa tgattttatt ttattcttct tcatataaat gtttaatatata 360
caatataaac aaattcttta ccttaagaag gatttcccat ttttatattt aaaaatatat 420
ttatcaaata tttttcaacc acgtaaactc cataataata agttgtttca aaagtaataa 480
aatttaactc cataattttt ttattcgact gatcttaag caacacccag tgacacaact 540
agccattttt ttctttgaat aaaaaaatcc aattatcatt gtattttttt tatacaatga 600
aaatttcacc aaacaatcat ttgtggtatt tctgaagcaa gtcagtgtat gcaaaattct 660
ataattccca ttgtacacta cggaagtaac tgaagatctg cttttacatg cgagacacat 720
cttctaaagt aatttttaata atagttacta tattcaagat ttcatatatc aaatactcaa 780
tattacttct aaaaaattaa ttagatataa ttaaaatatt acttttttaa ttttaagttt 840
aattgttgaa ttgttgacta ttgatttatt attctactat gtttaaattg ttttatagat 900
agtttaaagt aaatataagt aatgtagtag agtggttagag tgttacccta aaccataaac 960
tataagattt atggtggact aattttcata tatttcttat tgcttttacc ttttcttggg 1020
atgtaagtcc gtaactggaa ttactgtggg ttgccatggc actctgtggg cttttggttc 1080
atgcatggat gcttgcgcaa gaaaaagaca aagaacaaag aaaaaagaca aaacagagag 1140
acaaaacgca atcacacaac caactcaaat tagtactggg ctgatcaaga tcgcgcgctc 1200
catgtatgtc taaatgccat gcaaagcaac acgtgcttaa catgcacttt aaatggctca 1260
cccattctca cccacacaca aacacattgc ctttttcttc atcatcacca caaccacctg 1320
tatatattca ttctcttccg ccacctcaat ttcttcactt caacacacgt caacctgcat 1380
atgcgtgtca tcccatgccc aaatctccat gcatgttcca accaccttct ctcttatata 1440
atacctataa atacctctaa tatcactcac ttctttcatc atccatccat ccagagtact 1500
actactctac tactataata cccaaccca actcatattc aatactactc tact atg 1557
Met
1

```

```

gcg gat aca gct aga gga acc cat cac gat atc atc ggc aga gac cag 1605
Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp Gln
5 10 15

```

```

tac ccg atg atg ggc cga gac cga gac cag tac cag atg tcc gga cga 1653
Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly Arg
20 25 30

```

```

gga tct gac tac tcc aag tct agg cag att gct aaa gct gca act gct 1701
Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr Ala
35 40 45

```

```

gtc aca gct ggt ggt tcc ctc ctt gtt ctc tcc agc ctt acc ctt gtt 1749
Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu Val
50 55 60 65

```

```

gga act gtc ata gct ttg act gtt gca aca cct ctg ctc gtt atc ttc 1797
Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile Phe
70 75 80

```

```

agc cca atc ctt gtc ccg gct ctc atc aca gtt gca ctc ctc atc acc 1845
Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile Thr
85 90 95

```

```

ggt ttt ctt tcc tct gga ggg ttt ggc att gcc gct ata acc gtt ttc 1893
Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val Phe
100 105 110

```

```

tct tgg att tac aa gtaagcacac atttatcatc ttacttcata attttgtgca 1947
Ser Trp Ile Tyr Lys
115

```

```

atatgtgcat gcatgtgttg agccagtagc tttggatcaa tttttttggt cgaataacaa 2007
atgtaacaat aagaaattgc aaattctagg gaacatttgg ttaactaaat acgaaatttg 2067
acctagctag cttgaatgtg tctgtgtata tcatctatat aggtaaaatg cttggtatga 2127
tacctattga ttgtgaatag g tac gca acg gga gag cac cca cag gga tca 2178
Tyr Ala Thr Gly Glu His Pro Gln Gly Ser

```

	120	125	
gac aag ttg gac agt gca agg atg aag ttg gga agc aaa gct cag gat			2226
Asp Lys Leu Asp Ser Ala Arg Met Lys Leu Gly Ser Lys Ala Gln Asp			
130	135	140	
ctg aaa gac aga gct cag tac tac gga cag caa cat act ggt ggg gaa			2274
Leu Lys Asp Arg Ala Gln Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu			
145	150	155	160
cat gac cgt gac cgt act cgt ggt ggc cag cac act acc atg gct tcg			2322
His Asp Arg Asp Arg Thr Arg Gly Gly Gln His Thr Thr Met Ala Ser			
165	170	175	
gaa gaa gga caa gtg atc gcc tgc cac acc gtt gag aca tgg aac gag			2370
Glu Glu Gly Gln Val Ile Ala Cys His Thr Val Glu Thr Trp Asn Glu			
180	185	190	
cag ctt cag aag gct aat gaa tcc aaa act ctt gtg gtg gtt gat ttc			2418
Gln Leu Gln Lys Ala Asn Glu Ser Lys Thr Leu Val Val Val Asp Phe			
195	200	205	
acg gct tct tgg tgt gga cca tgt cgt ttc atc gct cca ttc ttt gct			2466
Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro Phe Phe Ala			
210	215	220	
gat ttg gct aag aaa ctt cct aac gtg ctt ttc ctc aag gtt gat act			2514
Asp Leu Ala Lys Lys Leu Pro Asn Val Leu Phe Leu Lys Val Asp Thr			
225	230	235	240
gat gaa ttg aag tcg gtg gca agt gat tgg gcg ata cag gcg atg cca			2562
Asp Glu Leu Lys Ser Val Ala Ser Asp Trp Ala Ile Gln Ala Met Pro			
245	250	255	
acc ttc atg ttt ttg aag gaa ggg aag att ttg gac aaa gtt gtt gga			2610
Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Leu Asp Lys Val Val Gly			
260	265	270	
gcc aag aaa gat gag ctt cag tct acc att gcc aaa cac ttg gct taa			2658
Ala Lys Lys Asp Glu Leu Gln Ser Thr Ile Ala Lys His Leu Ala *			
275	280	285	
g cttaataagt atgaactaaa atgcatgtag gtgtaagagc tcatggagag			2709
catggaatat tgtatccgac catgtaacag tataataact gagctccatc tcaacttcttc			2769
tatgaataaaa caaaggatgt tatgatatat taacactcta tctatgcacc ttattgttct			2829
atgataaatt tcctcttatt attataaatc atctgaatcg tgacggctta tggaaatgctt			2889
caaatagtag aaaaacaaat gtgtactata agactttcta aacaattcta acttttagcat			2949
tgtgaacgag acataagtgt taagaagaca taacaattat aatggaagaa gtttgtctcc			3009
atttatatat tatatattac ccacttatgt attatattag gatgttaagg agacataaca			3069
attataaaga gagaagtttg tatccattta tatattatat actaccatt tatatattat			3129
acttatccac ttatttaagtg tctttataag gtttgatcca tgatatttct aatattttag			3189
ttgatatgta tatgaaaggg tactatttga actctcttac tctgtataaa ggttggatca			3249
tccttaaagt gggctctattt aattttattg cttcttacag ataaaaaaaa aattatgagt			3309
tggtttgata aaatattgaa ggatttaaaa taataataaa taataataaa catataatat			3369
atgtatataa atttattata atataacatt tatctataaa aaagtaaata ttgtcataaa			3429
tctatacaat cgtttagcct tgctggacga ctctcaatta tttaaacgag agtaaacata			3489
tttgactttt tggttattta acaaaattatt atttaacact atatgaaatt tttttttttt			3549
atcggcaagg aaataaaatt aaattaggag ggacaatggt gtgtcccaat ccttatataa			3609
ccaacttcca caggaagggt aggtcgggga caacaaaaaa acaggcaagg gaaatttttt			3669
aatttgggtt gtcttgtttg ctgcataatt tatgcagtaa aacactacac ataacccttt			3729
tagcagtaga gcaatgggtg accgtgtgct tagcttcttt tatttttatt tttttatcagc			3789
aaagaataaa taaaataaaa tgagacactt cagggatgtt tcaaccctta tacaaaacc			3849
caaaaacaag tttcctagca ccctaccaac taagggtacc			3888

<210> 17

<211> 118

<212> PRT

<213> Artificial sequence

<400> 17

Met Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp
 1 5 10 15
 Gln Tyr Pro Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly
 20 25 30
 Arg Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr
 35 40 45
 Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu
 50 55 60
 Val Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile
 65 70 75 80
 Phe Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile
 85 90 95
 Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val
 100 105 110
 Phe Ser Trp Ile Tyr Lys
 115

<210> 18

<211> 169

<212> PRT

<213> Artificial sequence

<400> 18

Tyr Ala Thr Gly Glu His Pro Gln Gly Ser Asp Lys Leu Asp Ser Ala
 1 5 10 15
 Arg Met Lys Leu Gly Ser Lys Ala Gln Asp Leu Lys Asp Arg Ala Gln
 20 25 30
 Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu His Asp Arg Asp Arg Thr
 35 40 45
 Arg Gly Gly Gln His Thr Thr Met Ala Ser Glu Glu Gly Gln Val Ile
 50 55 60
 Ala Cys His Thr Val Glu Thr Trp Asn Glu Gln Leu Gln Lys Ala Asn
 65 70 75 80
 Glu Ser Lys Thr Leu Val Val Val Asp Phe Thr Ala Ser Trp Cys Gly
 85 90 95
 Pro Cys Arg Phe Ile Ala Pro Phe Phe Ala Asp Leu Ala Lys Lys Leu
 100 105 110
 Pro Asn Val Leu Phe Leu Lys Val Asp Thr Asp Glu Leu Lys Ser Val
 115 120 125
 Ala Ser Asp Trp Ala Ile Gln Ala Met Pro Thr Phe Met Phe Leu Lys
 130 135 140
 Glu Gly Lys Ile Leu Asp Lys Val Val Gly Ala Lys Lys Asp Glu Leu
 145 150 155 160
 Gln Ser Thr Ile Ala Lys His Leu Ala
 165

<210> 19

<211> 3888

<212> DNA

<213> Artificial Sequence

<220>

<223> Chimeric

<221> CDS

<222> (1555)...(2249)

<221> CDS

<222> (2490)...(2658)

<400> 19

ctgcaggaat tcattgtact cccagtatca ttatagttaa agttttggct ctctcgccgg 60

```

tgggtttttta cctctatttta aagggggtttt ccacctaataa attctggtat cattctcact 120
ttacttggtta ctttaattttc tcataatctt tgggtgaaat tatcacgctt ccgcacacga 180
tatccctaca aattttattat ttgttaaaca ttttcaaaac gcataaaatt ttatgaagtc 240
ccgtctatct ttaatgtagt ctaacattttt catattgaaa tatataattt acttaatttt 300
agcgttggtta gaaagcataa tgattttattc ttattcttct tcatataaat gtttaatatata 360
caatataaac aaattcttta ccttaagaag gatttcccat tttatatattt aaaaatatata 420
ttatcaaata tttttcaacc acgtaaatct cataataata agttggttca aaagtaataa 480
aatttaactc cataattttt ttattcgact gatcttaaag caacaccag tgacacaact 540
agccattttt ttctttgaaat aaaaaaatcc aattatcatt gtattttttt tatacaatga 600
aaatttcacc aaacaatcat ttgtggtatt tctgaagcaa gtcattgtat gcaaaattct 660
ataattccca ttgacacta cggaagtaac tgaagatctg cttttacatg cgagacacat 720
cttctaaggt aatttttaata atagttacta tattcaagat ttcatatata aaataactca 780
tattacttct aaaaaattaa ttagatataa ttaaaatatt acttttttaa ttttaagttt 840
aattgttgaa tttgtgacta ttgattttatt attctactat gtttaaattg ttttatagat 900
agtttaaagt aaatataagt aatgtagtag agtggttagag tgttacccta aaccataaac 960
tataagattt atggtggact aattttcata tatttcttat tgcttttacc ttttcttggt 1020
atgtaagttc gtaactggaa ttactgtggg ttgccatggc actctgtggt cttttgttgc 1080
atgcatggat gcttgcgcaa gaaaaagaca aagaacaaag aaaaaagaca aaacagagag 1140
acaaaacgca atcacacaac caactcaaat tagtctactg ctgatcaaga tcgccgcgtc 1200
catgtatgtc taaatgccat gcaaagcaac acgtgcttaa catgcacttt aaatggctca 1260
cccattctca cccacacaca aacacattgc ctttttcttc atcatcacca caaccactg 1320
tatattattca ttctcttcg ccacctcaat ttcttcaact caacacacgt caacctgcac 1380
atgcgtgtca tcccatgccc aaatctccat gcatgttcca accaccttct ctcttatata 1440
atactataaa atacctctaa tatcactcac ttcttctatc atccatccat ccagagtact 1500
actactctac tactataata cccaaccca actcatattc aatactactc tact atg 1557

```

Met
1

```

gct tgc gaa gaa gga caa gtg atc gcc tgc cac acc gtt gag aca tgg 1605
Ala Ser Glu Glu Gly Gln Val Ile Ala Cys His Thr Val Glu Thr Trp
          5                10                15

```

```

aac gag cag ctt cag aag gct aat gaa tcc aaa act ctt gtg gtg gtt 1653
Asn Glu Gln Leu Gln Lys Ala Asn Glu Ser Lys Thr Leu Val Val Val
          20                25                30

```

```

gat ttc acg gct tct tgg tgt gga cca tgt cgt ttc atc gct cca ttc 1701
Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro Phe
          35                40                45

```

```

ttt gct gat ttg gct aag aaa ctt cct aac gtg ctt ttc ctc aag gtt 1749
Phe Ala Asp Leu Ala Lys Lys Leu Pro Asn Val Leu Phe Leu Lys Val
          50                55                60                65

```

```

gat act gat gaa ttg aag tgc gtg gca agt gat tgg gcg ata cag gcg 1797
Asp Thr Asp Glu Leu Lys Ser Val Ala Ser Asp Trp Ala Ile Gln Ala
          70                75                80

```

```

atg cca acc ttc atg ttt ttg aag gaa ggg aag att ttg gac aaa gtt 1845
Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Leu Asp Lys Val
          85                90                95

```

```

gtt gga gcc aag aaa gat gag ctt cag tct acc att gcc aaa cac ttg 1893
Val Gly Ala Lys Lys Asp Glu Leu Gln Ser Thr Ile Ala Lys His Leu
          100                105                110

```

```

gct atg gcg gat aca gct aga gga acc cat cac gat atc atc ggc aga 1941
Ala Met Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg
          115                120                125

```

```

gac cag tac ccg atg atg ggc cga gac cga gac cag tac cag atg tcc 1989
Asp Gln Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser
          130                135                140                145

```

```

gga cga gga tct gac tac tcc aag tct agg cag att gct aaa gct gca 2037
Gly Arg Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala
          150                155                160

```

act gct gtc aca gct ggt ggt tcc ctc ctt gtt ctc tcc agc ctt acc 2085
 Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr
 165 170 175

ctt gtt gga act gtc ata gct ttg act gtt gca aca cct ctg ctc gtt 2133
 Leu Val Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val
 180 185 190

atc ttc agc cca atc ctt gtc ccg gct ctc atc aca gtt gca ctc ctc 2181
 Ile Phe Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu
 195 200 205

atc acc ggt ttt ctt tcc tct gga ggg ttt ggc att gcc gct ata acc 2229
 Ile Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr
 210 215 220 225

gtt ttc tct tgg att tac aa gtaagcacac atttatcatc ttacttcata 2279
 Val Phe Ser Trp Ile Tyr Lys
 230

attttgtgca atatgtgcat gcatgtgttg agccagtagc tttggatcaa tttttttggt 2339
 cgaataacaa atgtaacaat aagaaattgc aaattctagg gaacatttgg ttaactaaat 2399
 acgaaatttg acctagctag cttgaatgtg tctgtgtata tcatctatat aggtaaaaatg 2459
 cttggtatga tacctattga ttgtgaatag g tac gca acg gga gag cac cca 2511
 Tyr Ala Thr Gly Glu His Pro
 235

cag gga tca gac aag ttg gac agt gca agg atg aag ttg gga agc aaa 2559
 Gln Gly Ser Asp Lys Leu Asp Ser Ala Arg Met Lys Leu Gly Ser Lys
 240 245 250 255

gct cag gat ctg aaa gac aga gct cag tac tac gga cag caa cat act 2607
 Ala Gln Asp Leu Lys Asp Arg Ala Gln Tyr Tyr Gly Gln Gln His Thr
 260 265 270

ggt ggg gaa cat gac cgt gac cgt act cgt ggt ggc cag cac act act 2655
 Gly Gly Glu His Asp Arg Asp Arg Thr Arg Gly Gly Gln His Thr Thr
 275 280 285

taa gcttaataag tatgaactaa aatgcatgta ggtgtaagag ctcattggaga 2708
 *

gcatggaata ttgtatccga ccatgtaaca gtataataac tgagctccat ctcaacttctt 2768
 ctatgaataa acaaaggatg ttatgatata ttaacactct atctatgcac cttattgttc 2828
 tatgataaat ttcctcttat tattataaat catctgaatc gtgacggctt atggaatgct 2888
 tcaaatagta caaaaacaaa tgtgtactat aagactttct aaacaattct aacttttagca 2948
 ttgtgaacga gacataagtg ttaagaagac ataacaatta taatggaaga agtttgtctc 3008
 catttatata ttatatatta cccacttatg tattatatta ggatgttaag gagacataac 3068
 aattataaag agagaagttt gtatccattt atatattata tactaccat ttatatatta 3128
 tacttatcca cttattttaat gtctttataa ggtttgatcc atgatatttc taatatttta 3188
 gttgatatgt atatgaaagg gtactatttg aactctctta ctctgtataa aggttggatc 3248
 atccttaaag tgggtctatt taattttatt gcttcttaca gataaaaaaa aaattatgag 3308
 ttggtttgat aaaatattga aggatttaaa ataataataa ataataaata acatataata 3368
 tatgtatata aattttattat aatataacat ttatctataa aaaagtaaat attgtcataa 3428
 atctatacaa tcgttttagcc ttgctggacg actctcaatt atttaaacga gagtaaacat 3488
 atttgacttt ttggttattht aacaaattat tatttaaacac tatatgaaat tttttttttt 3548
 tatcggaag gaaataaaat taaattagga gggacaatgg tgtgtcccaa tccttatata 3608
 accaacttcc acaggaaggt caggctgggg acaacaaaaa aacaggcaag ggaaattttt 3668
 taatttgggt tgtcttggtt gctgcataat ttatgcagta aaacactaca cataaccctt 3728
 ttagcagtag agcaatgggt gaccgtgtgc tttagcttctt ttattttatt tttttatcag 3788
 caaagaataa ataaaaataa atgagacact tcagggatgt ttcaaccctt atacaaaacc 3848
 caaaaacaa gtttcctagc accctaccaa ctaaggtacc 3888

<210> 20
 <211> 232

<212> PRT
 <213> Artificial Sequence

<220>
 <223> Chimeric

<400> 20
 Met Ala Ser Glu Glu Gly Gln Val Ile Ala Cys His Thr Val Glu Thr
 1 5 10 15
 Trp Asn Glu Gln Leu Gln Lys Ala Asn Glu Ser Lys Thr Leu Val Val
 20 25 30
 Val Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro
 35 40 45
 Phe Phe Ala Asp Leu Ala Lys Lys Leu Pro Asn Val Leu Phe Leu Lys
 50 55 60
 Val Asp Thr Asp Glu Leu Lys Ser Val Ala Ser Asp Trp Ala Ile Gln
 65 70 75 80
 Ala Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Leu Asp Lys
 85 90 95
 Val Val Gly Ala Lys Lys Asp Glu Leu Gln Ser Thr Ile Ala Lys His
 100 105 110
 Leu Ala Met Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly
 115 120 125
 Arg Asp Gln Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met
 130 135 140
 Ser Gly Arg Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala
 145 150 155 160
 Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu
 165 170 175
 Thr Leu Val Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu
 180 185 190
 Val Ile Phe Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu
 195 200 205
 Leu Ile Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile
 210 215 220
 Thr Val Phe Ser Trp Ile Tyr Lys
 225 230

<210> 21
 <211> 55
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Chimeric

<400> 21
 Tyr Ala Thr Gly Glu His Pro Gln Gly Ser Asp Lys Leu Asp Ser Ala
 1 5 10 15
 Arg Met Lys Leu Gly Ser Lys Ala Gln Asp Leu Lys Asp Arg Ala Gln
 20 25 30
 Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu His Asp Arg Asp Arg Thr
 35 40 45
 Arg Gly Gly Gln His Thr Thr
 50 55

<210> 22
 <211> 3787
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Chimeric

<221> CDS

<222> (1555) ... (2556)

<400> 22

```

ctgcaggaat tcattgtact cccagtatca ttatagttaa agttttggct ctctcgccgg 60
tggtttttta cctctattta aagggggttt ccacctaaaa attctggtat cattctcact 120
ttactttgta ctttaatttc tcataatcct tggttgaaat tatcacgctt ccgcacacga 180
tatccctaca aattttattat ttgttaaaca ttttcaaacc gcataaaatt ttatgaagtc 240
ccgtctatct ttaatgtagt ctaacatttt catattgaaa tatataattt acttaatttt 300
agcgttggtg gaaagcataa tgattttatt ttattcttct tcatataaat gtttaatatata 360
caatataaac aaattcttta ccttaagaag gattttccat tttatatttt aaaaatatata 420
ttatcaaata tttttcaacc acgtaaactc cataataata agttgtttca aaagtaataa 480
aatttaactc cataattttt ttattcgact gatcttaaa caacacccag tgacacaact 540
agccattttt ttctttgaat aaaaaaatcc aattatcatt gtattttttt tatacaatga 600
aaatttcacc aaacaatcat ttgtggtatt tctgaagcaa gtcattgtat gcaaaattct 660
ataattccca tttgacacta cggaagtaac tgaagatctg cttttacatg cgagacacat 720
cttctaaagt aattttaata atagttacta tattcaagat ttcatatata aaataactca 780
tattacttct aaaaaattaa ttagatataa ttaaaatatt acttttttaa ttttaagttt 840
aattggtgaa tttgtgacta ttgatttatt attctactat gtttaaattg ttttatagat 900
agtttaaagt aaatataagt aatgtagtag agtggttagag tgttacccta aaccataaac 960
tataagattt atggtggact aattttcata tattttctat tgcttttacc ttttcttggg 1020
atgtaagtcc gtaactggaa ttactgtggg ttgccatggc actctgtggt cttttggttc 1080
atgcatggat gcttgcgcaa gaaaaagaca aagaacaaag aaaaaagaca aaacagagag 1140
acaaaacgca atcacacaac caactcaaat tagtctactg ctgatcaaga tcgcccgctc 1200
catgtatgtc taaatgccat gcaaagcaac acgtgcttaa catgcacttt aaatggctca 1260
cccattctca cccacacaca aacacattgc ctttttcttc atcatcacca caaccactg 1320
tatatttca ttctcttcg ccacctcaat ttcttcaact caacacacgt caacctgcat 1380
atgcgtgtca tcccatgccc aaatctccat gcatgttcca accaccttct ctcttatata 1440
atactataa atacctctaa tatcactcac ttctttcatc atccatccat ccagagtact 1500
actactctac tactataata cccaaccca actcatattc aatactactc tact atg 1557

```

Met
1

```

aat ggt ctc gaa act cac aac aca agg ctc tgt atc gta gga agt ggc 1605
Asn Gly Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser Gly
      5                10                15

```

```

cca gcg gca cac acg gcg gcg att tac gca gct agg gct gaa ctt aaa 1653
Pro Ala Ala His Thr Ala Ala Ile Tyr Ala Ala Arg Ala Glu Leu Lys
      20                25                30

```

```

cct ctt ctc ttc gaa gga tgg atg gct aac gac atc gct ccc ggt ggt 1701
Pro Leu Leu Phe Glu Gly Trp Met Ala Asn Asp Ile Ala Pro Gly Gly
      35                40                45

```

```

caa cta aca acc acc acc gac gtc gag aat ttc ccc gga ttt cca gaa 1749
Gln Leu Thr Thr Thr Asp Val Glu Asn Phe Pro Gly Phe Pro Glu
      50                55                60                65

```

```

ggt att ctc gga gta gag ctc act gac aaa ttc cgt aaa caa tcg gag 1797
Gly Ile Leu Gly Val Glu Leu Thr Asp Lys Phe Arg Lys Gln Ser Glu
      70                75                80

```

```

cga ttc ggt act acg ata ttt aca gag acg gtg acg aaa gtc gat ttc 1845
Arg Phe Gly Thr Thr Ile Phe Thr Glu Thr Val Thr Lys Val Asp Phe
      85                90                95

```

```

tct tcg aaa ccg ttt aag cta ttc aca gat tca aaa gcc att ctc gct 1893
Ser Ser Lys Pro Phe Lys Leu Phe Thr Asp Ser Lys Ala Ile Leu Ala
      100                105                110

```

```

gac gct gtg att ctc gct act gga gct gtg gct aag cgg ctt agc ttc 1941
Asp Ala Val Ile Leu Ala Thr Gly Ala Val Ala Lys Arg Leu Ser Phe
      115                120                125

```

```

gtt gga tct ggt gaa ggt tct gga ggt ttc tgg aac cgt gga atc tcc 1989
Val Gly Ser Gly Glu Gly Ser Gly Gly Phe Trp Asn Arg Gly Ile Ser
      130                135                140                145

```

gct tgt gct gtt tgc gac gga gct gct ccg ata ttc cgt aac aaa cct	2037
Ala Cys Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg Asn Lys Pro	
150 155 160	
ctt gcg gtg atc ggt gga ggc gat tca gca atg gaa gaa gca aac ttt	2085
Leu Ala Val Ile Gly Gly Gly Asp Ser Ala Met Glu Glu Ala Asn Phe	
165 170 175	
ctt aca aaa tat gga tct aaa gtg tat ata atc cat agg aga gat gct	2133
Leu Thr Lys Tyr Gly Ser Lys Val Tyr Ile Ile His Arg Arg Asp Ala	
180 185 190	
ttt aga gcg tct aag att atg cag cag cga gct ttg tct aat cct aag	2181
Phe Arg Ala Ser Lys Ile Met Gln Gln Arg Ala Leu Ser Asn Pro Lys	
195 200 205	
att gat gtg att tgg aac tcg tct gtt gtg gaa gct tat gga gat gga	2229
Ile Asp Val Ile Trp Asn Ser Ser Val Val Glu Ala Tyr Gly Asp Gly	
210 215 220 225	
gaa aga gat gtg ctt gga gga ttg aaa gtg aag aat gtg gtt acc gga	2277
Glu Arg Asp Val Leu Gly Gly Leu Lys Val Lys Asn Val Val Thr Gly	
230 235 240	
gat gtt tct gat tta aaa gtt tct gga ttg ttc ttt gct att ggt cat	2325
Asp Val Ser Asp Leu Lys Val Ser Gly Leu Phe Phe Ala Ile Gly His	
245 250 255	
gag cca gct acc aag ttt ttg gat ggt ggt gtt gag tta gat tcg gat	2373
Glu Pro Ala Thr Lys Phe Leu Asp Gly Gly Val Glu Leu Asp Ser Asp	
260 265 270	
ggt tat gtt gtc acg aag cct ggt act aca cag act agc gtt ccc gga	2421
Gly Tyr Val Val Thr Lys Pro Gly Thr Thr Gln Thr Ser Val Pro Gly	
275 280 285	
ggt ttc gct gcg ggt gat gtt cag gat aag aag tat agg caa gcc atc	2469
Val Phe Ala Ala Gly Asp Val Gln Asp Lys Lys Tyr Arg Gln Ala Ile	
290 295 300 305	
act gct gca gga act ggg tgc atg gca gct ttg gat gca gag cat tac	2517
Thr Ala Ala Gly Thr Gly Cys Met Ala Ala Leu Asp Ala Glu His Tyr	
310 315 320	
tta caa gag att gga tct cag caa ggt aag agt gat tga agcttaataa	2566
Leu Gln Glu Ile Gly Ser Gln Gln Gly Lys Ser Asp *	
325 330	
gtatgaacta aaatgcatgt aggtgtaaga gctcatggag agcatggaat attgtatccg	2626
accatgtaac agtataataa ctgagctcca tctcacttct tctatgaata aacaaaggat	2686
ggtatgatata attaacactc tatctatgca ccttattgtt ctatgataaa tttcctctta	2746
ttattataaa tcatctgaat cgtgacggct tatggaatgc ttcaaatagt acaaaaacaa	2806
atgtgtacta taagactttc taaacaattc taacttttagc attgtgaacg agacataagt	2866
gttaagaaga cataacaatt ataatggaag aagtttgtct ccatttatat attatatatt	2926
accacttat gtattatatt aggatgttaa ggagacataa caattataaa gagagaagtt	2986
tgtatccatt tataatattat atactaccca tttatatatt atacttatcc acttatttaa	3046
tgtctttata aggtttgatc catgatattt ctaatatatt agttgatatg tatatgaaag	3106
ggtactatatt gaactctctt actctgtata aagggttgat catccttaaa gtgggtctat	3166
ttaattttat tgcttcttac agataaaaaa aaaattatga gttggtttga taaaatattg	3226
aaggatttaa aataataata aataataaat aacatataat atatgtatat aaatttatta	3286
taatataaca tttatctata aaaaagtaaa tattgtcata aatctataca atcgttttagc	3346
cttgctggac gactctcaat tatttaaacg agagtaaaca tatttgactt tttggttatt	3406
taacaaatta ttatttaaca ctatatgaaa tttttttttt ttatcggcaa ggaaataaaa	3466
ttaaattagg agggacaatg gtgtgtccca atccttatac aaccaacttc cacaggaagg	3526
tcaggctggg gacaacaaaa aaacaggcaa gggaaatttt ttaatttggg ttgtcttgtt	3586
tgctgcataa tttatgcagt aaaacactac acataaccct tttagcagta gagcaatggg	3646

tgaccgtgtg cttagcttct tttattttat ttttttatca gcaaagaata aataaaataa 3706
aatgagacac ttcagggatg tttcaaccct tatacaaaaac cccaaaaaca agtttcctag 3766
cacctacca actaaggtag c 3787

<210> 23
<211> 333
<212> PRT
<213> Artificial Sequence

<220>
<223> Chimeric

<400> 23
Met Asn Gly Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser
1 5 10 15
Gly Pro Ala Ala His Thr Ala Ala Ile Tyr Ala Ala Arg Ala Glu Leu
20 25 30
Lys Pro Leu Leu Phe Glu Gly Trp Met Ala Asn Asp Ile Ala Pro Gly
35 40 45
Gly Gln Leu Thr Thr Thr Thr Asp Val Glu Asn Phe Pro Gly Phe Pro
50 55 60
Glu Gly Ile Leu Gly Val Glu Leu Thr Asp Lys Phe Arg Lys Gln Ser
65 70 75 80
Glu Arg Phe Gly Thr Thr Ile Phe Thr Glu Thr Val Thr Lys Val Asp
85 90 95
Phe Ser Ser Lys Pro Phe Lys Leu Phe Thr Asp Ser Lys Ala Ile Leu
100 105 110
Ala Asp Ala Val Ile Leu Ala Thr Gly Ala Val Ala Lys Arg Leu Ser
115 120 125
Phe Val Gly Ser Gly Glu Gly Ser Gly Gly Phe Trp Asn Arg Gly Ile
130 135 140
Ser Ala Cys Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg Asn Lys
145 150 155 160
Pro Leu Ala Val Ile Gly Gly Gly Asp Ser Ala Met Glu Glu Ala Asn
165 170 175
Phe Leu Thr Lys Tyr Gly Ser Lys Val Tyr Ile Ile His Arg Arg Asp
180 185 190
Ala Phe Arg Ala Ser Lys Ile Met Gln Gln Arg Ala Leu Ser Asn Pro
195 200 205
Lys Ile Asp Val Ile Trp Asn Ser Ser Val Val Glu Ala Tyr Gly Asp
210 215 220
Gly Glu Arg Asp Val Leu Gly Gly Leu Lys Val Lys Asn Val Val Thr
225 230 235 240
Gly Asp Val Ser Asp Leu Lys Val Ser Gly Leu Phe Phe Ala Ile Gly
245 250 255
His Glu Pro Ala Thr Lys Phe Leu Asp Gly Gly Val Glu Leu Asp Ser
260 265 270
Asp Gly Tyr Val Val Thr Lys Pro Gly Thr Thr Gln Thr Ser Val Pro
275 280 285
Gly Val Phe Ala Ala Gly Asp Val Gln Asp Lys Lys Tyr Arg Gln Ala
290 295 300
Ile Thr Ala Ala Gly Thr Gly Cys Met Ala Ala Leu Asp Ala Glu His
305 310 315 320
Tyr Leu Gln Glu Ile Gly Ser Gln Gln Gly Lys Ser Asp
325 330

<210> 24
<211> 4546
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1555)...(1907)
<221> CDS

<222> (2148) ... (3315)

<223> Chimeric

<400> 24

```

ctgcaggaat tcattgtact cccagtatca ttatagttaa agttttggct ctctcgccgg 60
tggtttttta cctctattta aaggggtttt ccacctaaaa attctggtat cattctcact 120
ttacttggtta ctttaatttc tcataatcct tgggttgaaat tatcacgctt ccgcacacga 180
tatccctaca aattttattat ttgttaaaca ttttcaaacc gcataaaatt ttatgaagtc 240
ccgtctatct ttaagttagt ctaacatttt catattgaaa tatataattt acttaatttt 300
agcgttggta gaaagcataa tgatttattc ttattcttct tcatataaat gtttaataata 360
caatataaac aaattcttta ccttaagaag gatttcccat tttatatttt aaaaatataat 420
ttatcaaata tttttcaacc acgtaaactt cataataata agttggttca aaagtaataa 480
aatttaactc cataattttt ttattcgact gatcttaag caacaccag tgacacaact 540
agccattttt ttctttgaat aaaaaaatcc aattatcatt gtattttttt tatacaatga 600
aaatttcacc aaacaatcat ttgtggtatt tctgaagcaa gtcattgtat gcaaaattct 660
ataattccca ttgtgacta cggaagtaac tgaagatctg cttttacatg cgagacacat 720
cttctaaagt aatttttaata atagttacta tattcaagat ttcataatc aaatactcaa 780
tattacttct aaaaaattaa ttagatataa ttaaaatatt acttttttaa ttttaagttt 840
aattgttgaa ttgtgacta ttgatttatt attctactat gtttaaatg ttttatagat 900
agtttaaagt aaatataagt aatgtagtag agtggttagag tgttacccta aaccataaac 960
tataagattt atggtggact aattttcata tatttcttat tgcttttacc ttttcttggg 1020
atgtaagtc gtaactggaa ttactgtggg ttgccatggc actctgtggg cttttggttc 1080
atgcatggat gcttgcgcaa gaaaaagaca aagaacaaag aaaaaagaca aaacagagag 1140
acaaaacgca atcacacaac caactcaaat tagtcaactg ctgatcaaga tcgcccgcgc 1200
catgtatgtc taaatgccat gcaaagcaac acgtgcttaa catgcacttt aaatggctca 1260
cccattctca cccacacaca aacacattgc ctttttcttc atcatcacca caaccactg 1320
tatatattca ttctcttccg ccacctcaat ttcttcactt caacacacgt caacctgcat 1380
atgctgtgca tcccatgccc aaatctccat gcatgttcca accaccttct ctcttatata 1440
atacctataa atacctctaa tatcactcac ttctttcatc atccatccat ccagagtact 1500
actactctac tactataata ccccaacca actcatattc aatactactc tact atg 1557

```

Met

1

```

gcg gat aca gct aga gga acc cat cac gat atc atc ggc aga gac cag 1605
Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp Gln
5 10 15

```

```

tac ccg atg atg ggc cga gac cga gac cag tac cag atg tcc gga cga 1653
Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly Arg
20 25 30

```

```

gga tct gac tac tcc aag tct agg cag att gct aaa gct gca act gct 1701
Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr Ala
35 40 45

```

```

gtc aca gct ggt ggt tcc ctc ctt gtt ctc tcc agc ctt acc ctt gtt 1749
Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu Val
50 55 60 65

```

```

gga act gtc ata gct ttg act gtt gca aca cct ctg ctc gtt atc ttc 1797
Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile Phe
70 75 80

```

```

agc cca atc ctt gtc ccg gct ctc atc aca gtt gca ctc ctc atc acc 1845
Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile Thr
85 90 95

```

```

ggt ttt ctt tcc tct gga ggg ttt ggc att gcc gct ata acc gtt ttc 1893
Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val Phe
100 105 110

```

```

tct tgg att tac aa gtaagcacac atttatcatc ttacttcata attttgtgca 1947
Ser Trp Ile Tyr Lys
115

```

```

atatgtgcat gcatgtgttg agccagtagc tttggatcaa tttttttggg cgaataacaa 2007

```

```

atgtaacaat aagaaattgc aaattctagg gaacatttgg ttaactaaat acgaaatttg 2067
acctagctag cttgaatgtg tctgtgtata tcatctatat aggtaaaatg cttggtatga 2127
tacctattga ttgtgaatag g tac gca acg gga gag cac cca cag gga tca 2178
      Tyr Ala Thr Gly Glu His Pro Gln Gly Ser
      120                               125

gac aag ttg gac agt gca agg atg aag ttg gga agc aaa gct cag gat 2226
Asp Lys Leu Asp Ser Ala Arg Met Lys Leu Gly Ser Lys Ala Gln Asp
      130                               135                               140

ctg aaa gac aga gct cag tac tac gga cag caa cat act ggt ggg gaa 2274
Leu Lys Asp Arg Ala Gln Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu
      145                               150                               155                               160

cat gac cgt gac cgt act cgt ggt ggc cag cac act acc atg aat ggt 2322
His Asp Arg Asp Arg Thr Arg Gly Gly Gln His Thr Thr Met Asn Gly
      165                               170                               175

ctc gaa act cac aac aca agg ctc tgt atc gta gga agt ggc cca gcg 2370
Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser Gly Pro Ala
      180                               185                               190

gca cac acg gcg gcg att tac gca gct agg gct gaa ctt aaa cct ctt 2418
Ala His Thr Ala Ala Ile Tyr Ala Ala Arg Ala Glu Leu Lys Pro Leu
      195                               200                               205

ctc ttc gaa gga tgg atg gct aac gac atc gct ccc ggt ggt caa cta 2466
Leu Phe Glu Gly Trp Met Ala Asn Asp Ile Ala Pro Gly Gly Gln Leu
      210                               215                               220

aca acc acc acc gac gtc gag aat ttc ccc gga ttt cca gaa ggt att 2514
Thr Thr Thr Thr Asp Val Glu Asn Phe Pro Gly Phe Pro Glu Gly Ile
      225                               230                               235                               240

ctc gga gta gag ctc act gac aaa ttc cgt aaa caa tcg gag cga ttc 2562
Leu Gly Val Glu Leu Thr Asp Lys Phe Arg Lys Gln Ser Glu Arg Phe
      245                               250                               255

ggt act acg ata ttt aca gag acg gtg acg aaa gtc gat ttc tct tcg 2610
Gly Thr Thr Ile Phe Thr Glu Thr Val Thr Lys Val Asp Phe Ser Ser
      260                               265                               270

aaa ccg ttt aag cta ttc aca gat tca aaa gcc att ctc gct gac gct 2658
Lys Pro Phe Lys Leu Phe Thr Asp Ser Lys Ala Ile Leu Ala Asp Ala
      275                               280                               285

gtg att ctc gct act gga gct gtg gct aag cgg ctt agc ttc gtt gga 2706
Val Ile Leu Ala Thr Gly Ala Val Ala Lys Arg Leu Ser Phe Val Gly
      290                               295                               300

tct ggt gaa ggt tct gga ggt ttc tgg aac cgt gga atc tcc gct tgt 2754
Ser Gly Glu Gly Ser Gly Gly Phe Trp Asn Arg Gly Ile Ser Ala Cys
      305                               310                               315                               320

gct gtt tgc gac gga gct gct ccg ata ttc cgt aac aaa cct ctt gcg 2802
Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg Asn Lys Pro Leu Ala
      325                               330                               335

gtg atc ggt gga ggc gat tca gca atg gaa gaa gca aac ttt ctt aca 2850
Val Ile Gly Gly Gly Asp Ser Ala Met Glu Glu Ala Asn Phe Leu Thr
      340                               345                               350

aaa tat gga tct aaa gtg tat ata atc cat agg aga gat gct ttt aga 2898
Lys Tyr Gly Ser Lys Val Tyr Ile Ile His Arg Arg Asp Ala Phe Arg
      355                               360                               365

gcg tct aag att atg cag cag cga gct ttg tct aat cct aag att gat 2946

```

Ala	Ser	Lys	Ile	Met	Gln	Gln	Arg	Ala	Leu	Ser	Asn	Pro	Lys	Ile	Asp	
370						375					380					
gtg	att	tgg	aac	tcg	tct	gtt	gtg	gaa	gct	tat	gga	gat	gga	gaa	aga	2994
Val	Ile	Trp	Asn	Ser	Ser	Val	Val	Glu	Ala	Tyr	Gly	Asp	Gly	Glu	Arg	
385					390					395					400	
gat	gtg	ctt	gga	gga	ttg	aaa	gtg	aag	aat	gtg	gtt	acc	gga	gat	gtt	3042
Asp	Val	Leu	Gly	Gly	Leu	Lys	Val	Lys	Asn	Val	Val	Thr	Gly	Asp	Val	
			405						410					415		
tct	gat	tta	aaa	gtt	tct	gga	ttg	ttc	ttt	gct	att	ggg	cat	gag	cca	3090
Ser	Asp	Leu	Lys	Val	Ser	Gly	Leu	Phe	Phe	Ala	Ile	Gly	His	Glu	Pro	
			420					425					430			
gct	acc	aag	ttt	ttg	gat	ggg	ggg	gtt	gag	tta	gat	tcg	gat	ggg	tat	3138
Ala	Thr	Lys	Phe	Leu	Asp	Gly	Gly	Val	Glu	Leu	Asp	Ser	Asp	Gly	Tyr	
		435				440						445				
gtt	gtc	acg	aag	cct	ggg	act	aca	cag	act	agc	gtt	ccc	gga	gtt	ttc	3186
Val	Val	Thr	Lys	Pro	Gly	Thr	Thr	Gln	Thr	Ser	Val	Pro	Gly	Val	Phe	
	450					455					460					
gct	gcg	ggg	gat	gtt	cag	gat	aag	aag	tat	agg	caa	gcc	atc	act	gct	3234
Ala	Ala	Gly	Asp	Val	Gln	Asp	Lys	Lys	Tyr	Arg	Gln	Ala	Ile	Thr	Ala	
465					470				475						480	
gca	gga	act	ggg	tgc	atg	gca	gct	ttg	gat	gca	gag	cat	tac	tta	caa	3282
Ala	Gly	Thr	Gly	Cys	Met	Ala	Ala	Leu	Asp	Ala	Glu	His	Tyr	Leu	Gln	
			485						490					495		
gag	att	gga	tct	cag	caa	ggg	aag	agt	gat	tga	agcttaataa	gtatgaacta				3335
Glu	Ile	Gly	Ser	Gln	Gln	Gly	Lys	Ser	Asp	*						
		500						505								
aaatgcatgt	aggtgtaaga	gctcatggag	agcatggaat	attgtatccg	accatgtaac											3395
agtataataa	ctgagctcca	tctcacttct	tctatgaata	aacaaaggat	gttatgatat											3455
attaacactc	tatctatgca	ccttattggt	ctatgataaa	tttcctctta	ttattataaa											3515
tcacttgaat	cgtgacggct	tatggaatgc	ttcaaatagt	acaaaaacaa	atgtgtacta											3575
taagactttc	taaacaattc	taacttttagc	attgtgaacg	agacataagt	gttaagaaga											3635
cataacaatt	ataatggaag	aagtttgcct	ccatttatat	attatatatt	accacttat											3695
gtattatatt	aggatgtaa	ggagacataa	caattataaa	gagagaagtt	tgtatccatt											3755
tatatattat	atactaccca	tttatatatt	atacttatcc	acttatttaa	tgtctttata											3815
aggtttgatc	catgatattt	ctaataattt	agttgatatg	tatatgaaag	gggtactattt											3875
gaactctctt	actctgtata	aaggttggat	catccttaaa	gtgggtctat	ttaattttat											3935
tgcttcttac	agataaaaaa	aaaattatga	gttggtttga	taaaatattg	aaggatttaa											3995
aataataata	aataataaat	aacatataat	atatgtatat	aaatttatta	taatataaca											4055
tttatctata	aaaaagtaaa	tattgtcata	aatctataca	atcgtttagc	cttgctggac											4115
gactctcaat	tatttaaacg	agagtaaaca	tatttgactt	tttggttatt	taacaaatta											4175
ttatttaaca	ctatatgaaa	tttttttttt	ttatcggcaa	ggaaataaaa	ttaaattagg											4235
agggacaatg	gtgtgtccca	atccttatac	aaccaacttc	cacaggaagg	tcaggtcggg											4295
gacaacaaaa	aaacaggcaa	gggaaatttt	ttaatttggg	ttgtcttggt	tgctgcataa											4355
tttatgcagt	aaaacactac	acataaccct	tttagcagta	gagcaatggg	tgaccgtgtg											4415
cttagcttct	tttattttat	ttttttatca	gcaaagaata	aataaaataa	aatgagacac											4475
ttcagggatg	tttcaaccct	tatacaaaac	cccaaaaaac	agtttcctag	caccctacca											4535
actaaggtag	c															4546

<210> 25
 <211> 118
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Chimeric

<400> 25
 Met Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp

```

1      5      10      15
Gln Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly
20      25      30
Arg Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr
35      40      45
Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu
50      55      60
Val Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile
65      70      75      80
Phe Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile
85      90      95
Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val
100      105      110
Phe Ser Trp Ile Tyr Lys
115

```

<210> 26
 <211> 388
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Chimeric

```

<400> 26
Tyr Ala Thr Gly Glu His Pro Gln Gly Ser Asp Lys Leu Asp Ser Ala
1      5      10      15
Arg Met Lys Leu Gly Ser Lys Ala Gln Asp Leu Lys Asp Arg Ala Gln
20      25      30
Tyr Tyr Gly Gln Gln His Thr Gly Glu His Asp Arg Asp Arg Thr
35      40      45
Arg Gly Gly Gln His Thr Thr Met Asn Gly Leu Glu Thr His Asn Thr
50      55      60
Arg Leu Cys Ile Val Gly Ser Gly Pro Ala Ala His Thr Ala Ala Ile
65      70      75      80
Tyr Ala Ala Arg Ala Glu Leu Lys Pro Leu Leu Phe Glu Gly Trp Met
85      90      95
Ala Asn Asp Ile Ala Pro Gly Gly Gln Leu Thr Thr Thr Thr Asp Val
100      105      110
Glu Asn Phe Pro Gly Phe Pro Glu Gly Ile Leu Gly Val Glu Leu Thr
115      120      125
Asp Lys Phe Arg Lys Gln Ser Glu Arg Phe Gly Thr Thr Ile Phe Thr
130      135      140
Glu Thr Val Thr Lys Val Asp Phe Ser Ser Lys Pro Phe Lys Leu Phe
145      150      155      160
Thr Asp Ser Lys Ala Ile Leu Ala Asp Ala Val Ile Leu Ala Thr Gly
165      170      175      180
Ala Val Ala Lys Arg Leu Ser Phe Val Gly Ser Gly Glu Gly Ser Gly
180      185      190      195
Gly Phe Trp Asn Arg Gly Ile Ser Ala Cys Ala Val Cys Asp Gly Ala
195      200      205
Ala Pro Ile Phe Arg Asn Lys Pro Leu Ala Val Ile Gly Gly Gly Asp
210      215      220
Ser Ala Met Glu Glu Ala Asn Phe Leu Thr Lys Tyr Gly Ser Lys Val
225      230      235      240
Tyr Ile Ile His Arg Arg Asp Ala Phe Arg Ala Ser Lys Ile Met Gln
245      250      255
Gln Arg Ala Leu Ser Asn Pro Lys Ile Asp Val Ile Trp Asn Ser Ser
260      265      270
Val Val Glu Ala Tyr Gly Asp Gly Glu Arg Asp Val Leu Gly Gly Leu
275      280      285
Lys Val Lys Asn Val Val Thr Gly Asp Val Ser Asp Leu Lys Val Ser
290      295      300
Gly Leu Phe Phe Ala Ile Gly His Glu Pro Ala Thr Lys Phe Leu Asp
305      310      315      320
Gly Gly Val Glu Leu Asp Ser Asp Gly Tyr Val Val Thr Lys Pro Gly

```



```

          325          330          335
Thr Thr Gln Thr Ser Val Pro Gly Val Phe Ala Ala Gly Asp Val Gln
          340          345          350
Asp Lys Lys Tyr Arg Gln Ala Ile Thr Ala Ala Gly Thr Gly Cys Met
          355          360          365
Ala Ala Leu Asp Ala Glu His Tyr Leu Gln Glu Ile Gly Ser Gln Gln
          370          375          380
Gly Lys Ser Asp
385

```

<210> 27
 <211> 4545
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Chimeric

<221> CDS
 <222> (1555) ... (2906)

<221> CDS
 <222> (3147) ... (3315)

```

<400> 27
ctgcaggaat tcattgtact cccagtatca ttatagtga agttttggct ctctcgccgg 60
tggtttttta cctctattta aaggggtttt ccacctaaaa attctggtat cattctcact 120
ttacttggtta ctttaatttc tcataatctt tggttgaaat tatcacgctt ccgcacacga 180
tatccctaca aattttattat ttgttaaaca ttttcaaacc gcataaaatt ttatgaagtc 240
ccgtctatct ttaatgtagt ctaacatttt catattgaaa tatataattt acttaatttt 300
agcgttggta gaaagcataa tgattttatt ttattcttct tcatataaat gtttaatata 360
caatataaac aaattcttta ccttaagaag gatttcccat tttatatattt aaaaatata 420
ttatcaaata tttttcaacc acgtaaactt cataataata agttggttca aaagtaataa 480
aatttaacte cataattttt ttattcgact gatcttaaag caacacccag tgacacaact 540
agccattttt ttctttgaat aaaaaaatcc aattatcatt gtattttttt tatacaatga 600
aaatttcacc aaacaatcat ttgtggtatt tctgaagcaa gtcattgttat gcaaaattct 660
ataattccca ttgacacta cggaagtaac tgaagatctg cttttacatg cgagacacat 720
cttctaaagt aatttttaata atagttacta tattcaagat ttcatatata aaataactca 780
tattacttct aaaaaattaa ttagatataa ttaaaatatt acttttttaa ttttaagttt 840
aattgttgaa tttgtgacta ttgatttatt attctactat gtttaaattg ttttatagat 900
agtttaaggt aaatataagt aatgtagtag agtggttagag tgttacccta aaccataaac 960
tataagattt atgggtggact aattttcata tatttcttat tgcttttacc ttttcttgg 1020
atgtaagtc gtaactggaa ttactgtggg ttgccatggc actctgtggt cttttgggtc 1080
atgcatggat gcttgcgcaa gaaaaagaca aagaacaaag aaaaaagaca aaacagagag 1140
acaaaacgca atcacacaac caactcaaat tagtactgg ctgatcaaga tcgccgcgtc 1200
catgtatgtc taaatgccat gcaaaagcaac acgtgcttaa catgcacttt aaatggctca 1260
cccattctca cccacacaca aacacattgc ctttttcttc atcatcacca caaccacotg 1320
tatatattca ttctcttcg ccacctcaat ttcttcaact caacacacgt caacctgcat 1380
atgctgtgca tcccatgccc aaatctccat gcatgttcca accaccttct ctcttatata 1440
atacctataa atacctctaa tatcactcac ttctttcatc atccatccat ccagagtact 1500
actactctac tactataata ccccaaccca actcatattc aatactactc tact atg 1557
Met
1

```

```

aat ggt ctc gaa act cac aac aca agg ctc tgt atc gta gga agt ggc 1605
Asn Gly Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser Gly
          5          10          15

```

```

cca gcg gca cac acg gcg gcg att tac gca gct agg gct gaa ctt aaa 1653
Pro Ala Ala His Thr Ala Ala Ile Tyr Ala Ala Arg Ala Glu Leu Lys
          20          25          30

```

```

cct ctt ctc ttc gaa gga tgg atg gct aac gac atc gct ccc ggt ggt 1701
Pro Leu Leu Phe Glu Gly Trp Met Ala Asn Asp Ile Ala Pro Gly Gly
          35          40          45

```

caa cta aca acc acc acc gac gtc gag aat ttc ccc gga ttt cca gaa	1749
Gln Leu Thr Thr Thr Thr Asp Val Glu Asn Phe Pro Gly Phe Pro Glu	
50 55 60 65	
ggt att ctc gga gta gag ctc act gac aaa ttc cgt aaa caa tcg gag	1797
Gly Ile Leu Gly Val Glu Leu Thr Asp Lys Phe Arg Lys Gln Ser Glu	
70 75 80	
cga ttc ggt act acg ata ttt aca gag acg gtg acg aaa gtc gat ttc	1845
Arg Phe Gly Thr Thr Ile Phe Thr Glu Thr Val Thr Lys Val Asp Phe	
85 90 95	
tct tcg aaa ccg ttt aag cta ttc aca gat tca aaa gcc att ctc gct	1893
Ser Ser Lys Pro Phe Lys Leu Phe Thr Asp Ser Lys Ala Ile Leu Ala	
100 105 110	
gac gct gtg att ctc gct act gga gct gtg gct aag cgg ctt agc ttc	1941
Asp Ala Val Ile Leu Ala Thr Gly Ala Val Ala Lys Arg Leu Ser Phe	
115 120 125	
gtt gga tct ggt gaa ggt tct gga ggt ttc tgg aac cgt gga atc tcc	1989
Val Gly Ser Gly Glu Gly Ser Gly Gly Phe Trp Asn Arg Gly Ile Ser	
130 135 140 145	
gct tgt gct gtt tgc gac gga gct gct ccg ata ttc cgt aac aaa cct	2037
Ala Cys Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg Asn Lys Pro	
150 155 160	
ctt gcg gtg atc ggt gga ggc gat tca gca atg gaa gaa gca aac ttt	2085
Leu Ala Val Ile Gly Gly Gly Asp Ser Ala Met Glu Glu Ala Asn Phe	
165 170 175	
ctt aca aaa tat gga tct aaa gtg tat ata atc cat agg aga gat gct	2133
Leu Thr Lys Tyr Gly Ser Lys Val Tyr Ile Ile His Arg Arg Asp Ala	
180 185 190	
ttt aga gcg tct aag att atg cag cag cga gct ttg tct aat cct aag	2181
Phe Arg Ala Ser Lys Ile Met Gln Gln Arg Ala Leu Ser Asn Pro Lys	
195 200 205	
att gat gtg att tgg aac tcg tct gtt gtg gaa gct tat gga gat gga	2229
Ile Asp Val Ile Trp Asn Ser Ser Val Val Glu Ala Tyr Gly Asp Gly	
210 215 220 225	
gaa aga gat gtg ctt gga gga ttg aaa gtg aag aat gtg gtt acc gga	2277
Glu Arg Asp Val Leu Gly Gly Leu Lys Val Lys Asn Val Val Thr Gly	
230 235 240	
gat gtt tct gat tta aaa gtt tct gga ttg ttc ttt gct att ggt cat	2325
Asp Val Ser Asp Leu Lys Val Ser Gly Leu Phe Phe Ala Ile Gly His	
245 250 255	
gag cca gct acc aag ttt ttg gat ggt ggt gtt gag tta gat tcg gat	2373
Glu Pro Ala Thr Lys Phe Leu Asp Gly Gly Val Glu Leu Asp Ser Asp	
260 265 270	
ggt tat gtt gtc acg aag cct ggt act aca cag act agc gtt ccc gga	2421
Gly Tyr Val Val Thr Lys Pro Gly Thr Thr Gln Thr Ser Val Pro Gly	
275 280 285	
gtt ttc gct gcg ggt gat gtt cag gat aag aag tat agg caa gcc atc	2469
Val Phe Ala Ala Gly Asp Val Gln Asp Lys Lys Tyr Arg Gln Ala Ile	
290 295 300 305	
act gct gca gga act ggg tgc atg gca gct ttg gat gca gag cat tac	2517
Thr Ala Ala Gly Thr Gly Cys Met Ala Ala Leu Asp Ala Glu His Tyr	
310 315 320	

tta caa gag att gga tct cag caa ggt aag agt gat atg gcg gat aca 2565
 Leu Gln Glu Ile Gly Ser Gln Gln Gly Lys Ser Asp Met Ala Asp Thr
 325 330 335

gct aga gga acc cat cac gat atc atc ggc aga gac cag tac ccg atg 2613
 Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp Gln Tyr Pro Met
 340 345 350

atg ggc cga gac cga gac cag tac cag atg tcc gga cga gga tct gac 2661
 Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly Arg Gly Ser Asp
 355 360 365

tac tcc aag tct agg cag att gct aaa gct gca act gct gtc aca gct 2709
 Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr Ala Val Thr Ala
 370 375 380 385

ggt ggt tcc ctc ctt gtt ctc tcc agc ctt acc ctt gtt gga act gtc 2757
 Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu Val Gly Thr Val
 390 395 400

ata gct ttg act gtt gca aca cct ctg ctc gtt atc ttc agc cca atc 2805
 Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Ile
 405 410 415

ctt gtc ccg gct ctc atc aca gtt gca ctc ctc atc acc ggt ttt ctt 2853
 Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile Thr Gly Phe Leu
 420 425 430

tcc tct gga ggg ttt ggc att gcc gct ata acc gtt ttc tct tgg att 2901
 Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val Phe Ser Trp Ile
 435 440 445

tac aa gtaagcacac atttatcatc ttacttcata attttgtgca atatgtgcat 2956
 Tyr Lys
 450

gcatgtgttg agccagtagc tttggatcaa tttttttggt cgaataacaa atgtaacaat 3016
 aagaaattgc aaattctagg gaacatttgg ttaactaaat acgaaatttg acctagctag 3076
 cttgaatgtg tctgtgtata tcatctatat aggtaaaatg cttggtatga tacctattga 3136
 ttgtgaatag g tac gca acg gga gag cac cca cag gga tca gac aag ttg 3186
 Tyr Ala Thr Gly Glu His Pro Gln Gly Ser Asp Lys Leu
 455 460

gac agt gca agg atg aag ttg gga agc aaa gct cag gat ctg aaa gac 3234
 Asp Ser Ala Arg Met Lys Leu Gly Ser Lys Ala Gln Asp Leu Lys Asp
 465 470 475 480

aga gct cag tac tac gga cag caa cat act ggt ggg gaa cat gac cgt 3282
 Arg Ala Gln Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu His Asp Arg
 485 490 495

gac cgt act cgt ggt ggc cag cac act act taa gcttaataag tatgaactaa 3335
 Asp Arg Thr Arg Gly Gly Gln His Thr Thr *
 500 505

aatgcatgta ggtgtaagag ctcatggaga gcatggaata ttgtatccga ccatgtaaca 3395
 gtataataac tgagctccat ctcaatttct ctatgaataa acaaaggatg ttatgatata 3455
 ttaacactct atctatgcac ctattgttgc tatgataaat ttcctcttat tattataaat 3515
 catctgaatc gtgacggcct atggaatgct tcaaatagta caaaaacaaa tgtgtactat 3575
 aagactttct aaacaattct aactttagca ttgtgaacga gacataagtg ttaagaagac 3635
 ataacaatta taatggaaga agtttgtctc catttatata ttatatatta cccacttatg 3695
 tattatatta ggatgttaag gagacataac aattataaag agagaagttt gtatccattt 3755
 atatattata tactacccat ttatatatta tacttatcca cttatttaaat gtctttataa 3815
 ggtttgatcc atgatatttc taatatttta gttgatatgt atatgaaagg gtactatttg 3875
 aactctctta ctctgtataa aggttggatc atccttaaag tgggtctatt taattttatt 3935
 gcttcttaca gataaaaaaa aaattatgag ttggtttgat aaaatattga aggattttaa 3995

```

ataataataa ataataaata acatataata tatgtatata aatttattat aatataacat 4055
ttatctataa aaaagtaaat attgtcataa atctatacaa tcgttttagcc ttgctggacg 4115
actctcaatt atttaaacga gagtaaacat atttgacttt ttgggttattt aacaaattat 4175
tatttaacac tatatgaaat tttttttttt tatcggaag gaaataaaaat taaattagga 4235
gggacaatgg tgtgtcccaa tcctttatata accaacttcc acaggaaggt caggtcgggg 4295
acaacaaaaa aacaggcaag ggaaattttt taatttgggt tgtcttggtt gctgcataat 4355
ttatgcagta aaacactaca cataaccctt ttagcagtag agcaatgggt gaccgtgtgc 4415
ttagcttctt ttattttatt tttttatcag caaagaataa ataaaataaa atgagacact 4475
tcagggatgt ttcaaccctt atacaaaacc ccaaaaacaa gtttcctagc accctacaa 4535
ctaaggtacc 4545

```

<210> 28

<211> 451

<212> PRT

<213> Artificial Sequence

<220>

<223> Chimeric

<400> 28

```

Met Asn Gly Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser
 1          5          10          15
Gly Pro Ala Ala His Thr Ala Ala Ile Tyr Ala Ala Arg Ala Glu Leu
          20          25          30
Lys Pro Leu Leu Phe Glu Gly Trp Met Ala Asn Asp Ile Ala Pro Gly
          35          40          45
Gly Gln Leu Thr Thr Thr Thr Asp Val Glu Asn Phe Pro Gly Phe Pro
          50          55          60
Glu Gly Ile Leu Gly Val Glu Leu Thr Asp Lys Phe Arg Lys Gln Ser
65          70          75          80
Glu Arg Phe Gly Thr Thr Ile Phe Thr Glu Thr Val Thr Lys Val Asp
          85          90          95
Phe Ser Ser Lys Pro Phe Lys Leu Phe Thr Asp Ser Lys Ala Ile Leu
          100          105          110
Ala Asp Ala Val Ile Leu Ala Thr Gly Ala Val Ala Lys Arg Leu Ser
          115          120          125
Phe Val Gly Ser Gly Glu Gly Ser Gly Gly Phe Trp Asn Arg Gly Ile
          130          135          140
Ser Ala Cys Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg Asn Lys
145          150          155          160
Pro Leu Ala Val Ile Gly Gly Gly Asp Ser Ala Met Glu Glu Ala Asn
          165          170          175
Phe Leu Thr Lys Tyr Gly Ser Lys Val Tyr Ile Ile His Arg Arg Asp
          180          185          190
Ala Phe Arg Ala Ser Lys Ile Met Gln Gln Arg Ala Leu Ser Asn Pro
          195          200          205
Lys Ile Asp Val Ile Trp Asn Ser Ser Val Val Glu Ala Tyr Gly Asp
          210          215          220
Gly Glu Arg Asp Val Leu Gly Gly Leu Lys Val Lys Asn Val Val Thr
225          230          235          240
Gly Asp Val Ser Asp Leu Lys Val Ser Gly Leu Phe Phe Ala Ile Gly
          245          250          255
His Glu Pro Ala Thr Lys Phe Leu Asp Gly Gly Val Glu Leu Asp Ser
          260          265          270
Asp Gly Tyr Val Val Thr Lys Pro Gly Thr Thr Gln Thr Ser Val Pro
          275          280          285
Gly Val Phe Ala Ala Gly Asp Val Gln Asp Lys Lys Tyr Arg Gln Ala
          290          295          300
Ile Thr Ala Ala Gly Thr Gly Cys Met Ala Ala Leu Asp Ala Glu His
305          310          315          320
Tyr Leu Gln Glu Ile Gly Ser Gln Gln Gly Lys Ser Asp Met Ala Asp
          325          330          335
Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp Gln Tyr Pro
          340          345          350
Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly Arg Gly Ser
          355          360          365
Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr Ala Val Thr

```

```

      370              375              380
Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu Val Gly Thr
385              390              395              400
Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile Phe Ser Pro
      405              410              415
Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile Thr Gly Phe
      420              425              430
Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val Phe Ser Trp
      435              440              445
Ile Tyr Lys
      450

```

<210> 29
 <211> 55
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Chimeric

```

<400> 29
Tyr Ala Thr Gly Glu His Pro Gln Gly Ser Asp Lys Leu Asp Ser Ala
 1              5              10              15
Arg Met Lys Leu Gly Ser Lys Ala Gln Asp Leu Lys Asp Arg Ala Gln
      20              25              30
Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu His Asp Arg Asp Arg Thr
      35              40              45
Arg Gly Gly Gln His Thr Thr
      50              55

```

<210> 30
 <211> 4922
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> CDS
 <222> (1555) ... (1907)

<221> CDS
 <222> (2148) ... (3690)

<223> Chimeric

```

<400> 30
ctgcaggaat tcattgtact cccagtatca ttatagtgaag agttttggct ctctcgccgg 60
tggtttttta cctctattta aagggggttt ccacctaaaa attctggtat cattctcact 120
ttacttggtta ctttaatttc tcataatcct tgggtgaaat tatcacgctt ccgcacacga 180
tatccctaca aatttattat ttgttaaaca ttttcaaacc gcataaaatt ttatgaagtc 240
ccgtctatct ttaatgtagt ctaacatttt catattgaaa tatataattt acttaatttt 300
agcgttggtta gaaagcataa tgatttattc ttattcttct tcatataaat gtttaatatata 360
caatataaac aaattcttta ccttaagaag gatttcccat tttatatttt aaaaatatata 420
ttatcaaata tttttcaacc acgtaaactc cataataata agttgtttca aaagtaataa 480
aatttaactc cataattttt ttattcgact gatcttaaag caacaccag tgacacaact 540
agccattttt ttctttgaat aaaaaaatcc aattatcatt gtattttttt tatacaatga 600
aaatttcacc aaacaatcat ttgtggtatt tctgaagcaa gtcattgttat gcaaaattct 660
ataattccca tttgacacta cggaagtaac tgaagatctg cttttacatg cgagacacat 720
cttctaaagt aattttaata atagttacta tattcaagat ttcataatc aaataactca 780
tattacttct aaaaaattaa ttgatataaa ttaaaatatt acttttttaa ttttaagttt 840
aattgttgaa tttgtgacta ttgatttatt attctactat gtttaaattg ttttatagat 900
agtttaaagt aaatataagt aatgtagtag agtggttagag tgttacccta aaccataaac 960
tataagattt atggtggact aattttcata tatttcttat tgcttttacc ttttcttggt 1020
atgtaagtcc gtaactggaa ttactgtggg ttgccatggc actctgtggt cttttggttc 1080
atgcatggat gcttgcgcaa gaaaaagaca aaaaagaca aaacagagag 1140
acaaaacgca atcacacaac caactcaaat tagtcactgg ctgatcaaga tcgccgcgtc 1200

```

```

catgtatgtc taaatgccat gcaaagcaac acgtgcttaa catgcacttt aaatggctca 1260
cccatctcaa cccacacaca aacacattgc ctttttcttc atcatcacca caaccacctg 1320
tatatattca ttctcttccg ccacctcaat ttcttcactt caacacacgt caacctgcat 1380
atgcgtgtca tcccatgccc aaatctccat gcatgttcca accaccttct ctcttatata 1440
atacctataa atacctctaa tatcactcac ttctttcatc atccatccat ccagagtact 1500
actactctac tactataata ccccaaccga actcatattc aatactactc tact atg 1557
                                         Met
                                         1

!

gcg gat aca gct aga gga acc cat cac gat atc atc ggc aga gac cag 1605
Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp Gln
                                         5
                                         10
                                         15

tac ccg atg atg ggc cga gac cga gac cag tac cag atg tcc gga cga 1653
Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly Arg
                                         20
                                         25
                                         30

gga tct gac tac tcc aag tct agg cag att gct aaa gct gca act gct 1701
Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr Ala
                                         35
                                         40
                                         45

gtc aca gct ggt ggt tcc ctc ctt gtt ctc tcc agc ctt acc ctt gtt 1749
Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu Val
                                         50
                                         55
                                         60
                                         65

gga act gtc ata gct ttg act gtt gca aca cct ctg ctc gtt atc ttc 1797
Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile Phe
                                         70
                                         75
                                         80

agc cca atc ctt gtc ccg gct ctc atc aca gtt gca ctc ctc atc acc 1845
Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile Thr
                                         85
                                         90
                                         95

ggt ttt ctt tcc tct gga ggg ttt ggc att gcc gct ata acc gtt ttc 1893
Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val Phe
                                         100
                                         105
                                         110

tct tgg att tac aa gtaagcacac atttatcatc ttacttcata attttgtgca 1947
Ser Trp Ile Tyr Lys
                                         115

atatgtgcat gcatgtgttg agccagtagc tttggatcaa tttttttggt cgaataacaa 2007
atgtaacaat aagaaattgc aaattctagg gaacatttgg ttaactaaat acgaaatttg 2067
acctagctag cttgaatgtg tctgtgtata tcacttatat aggtaaaatg cttgggtatga 2127
tacctattga ttgtgaatag g tac gca acg gga gag cac cca cag gga tca 2178
                                         Tyr Ala Thr Gly Glu His Pro Gln Gly Ser
                                         120
                                         125

gac aag ttg gac agt gca agg atg aag ttg gga agc aaa gct cag gat 2226
Asp Lys Leu Asp Ser Ala Arg Met Lys Leu Gly Ser Lys Ala Gln Asp
                                         130
                                         135
                                         140

ctg aaa gac aga gct cag tac tac gga cag caa cat act ggt ggg gaa 2274
Leu Lys Asp Arg Ala Gln Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu
                                         145
                                         150
                                         155
                                         160

cat gac cgt gac cgt act cgt ggt ggc cag cac act acc atg aac acc 2322
His Asp Arg Asp Arg Thr Arg Gly Gly Gln His Thr Thr Met Asn Thr
                                         165
                                         170
                                         175

act cct tct gcg cat gag acg ata cac gaa gtg atc gtt att ggc tcc 2370
Thr Pro Ser Ala His Glu Thr Ile His Glu Val Ile Val Ile Gly Ser
                                         180
                                         185
                                         190

ggt cca gca ggc tac act gct gcc ctg tac gcc gct cgt gca cag cta 2418
Gly Pro Ala Gly Tyr Thr Ala Ala Leu Tyr Ala Ala Arg Ala Gln Leu
                                         195
                                         200
                                         205

```

aca ccg ctg gta ttt gag ggt acc tca ttc ggc ggc gcg ctg atg acc	2466
Thr Pro Leu Val Phe Glu Gly Thr Ser Phe Gly Gly Ala Leu Met Thr	
210 215 220	
acc acc gag gtg gaa aac tac cca ggt ttt cgc aac ggc ata acc ggc	2514
Thr Thr Glu Val Glu Asn Tyr Pro Gly Phe Arg Asn Gly Ile Thr Gly	
225 230 235 240	
ccg gag ttg atg gac gat atg cgt gaa cag gca ctg cga ttc ggc gcg	2562
Pro Glu Leu Met Asp Asp Met Arg Glu Gln Ala Leu Arg Phe Gly Ala	
245 250 255	
gaa ctg cgg acc gaa gac gtc gag tcg gta tca ttg cgt ggc ccg atc	2610
Glu Leu Arg Thr Glu Asp Val Glu Ser Val Ser Leu Arg Gly Pro Ile	
260 265 270	
aaa tcg gtc gtc acc gct gaa gga cag act tat cag gcc cga gcc gtc	2658
Lys Ser Val Val Thr Ala Glu Gly Gln Thr Tyr Gln Ala Arg Ala Val	
275 280 285	
atc ctc gcc atg ggt acc tcc gtg cgt tat cta cag atc ccc ggc gag	2706
Ile Leu Ala Met Gly Thr Ser Val Arg Tyr Leu Gln Ile Pro Gly Glu	
290 295 300	
caa gaa ttg cta gga cgt ggc gtg agt gca tgc gcg acc tgc gac ggg	2754
Gln Glu Leu Leu Gly Arg Gly Val Ser Ala Cys Ala Thr Cys Asp Gly	
305 310 315 320	
tcc ttt ttc cgc ggc caa gac att gcc gtc att ggc ggt gga gac tca	2802
Ser Phe Phe Arg Gly Gln Asp Ile Ala Val Ile Gly Gly Gly Asp Ser	
325 330 335	
gcg atg gag gaa gcc ctc ttt ttg acc cgg ttc gcc cgc agc gtc acg	2850
Ala Met Glu Glu Ala Leu Phe Leu Thr Arg Phe Ala Arg Ser Val Thr	
340 345 350	
ctc gtg cac cgc cgc gac gaa ttc cga gct tct aag atc atg ctc ggt	2898
Leu Val His Arg Arg Asp Glu Phe Arg Ala Ser Lys Ile Met Leu Gly	
355 360 365	
cgc gcc cgt aac aat gac aag atc aaa ttc atc acc aac cac acc gtg	2946
Arg Ala Arg Asn Asn Asp Lys Ile Lys Phe Ile Thr Asn His Thr Val	
370 375 380	
gtc gcg gtg aac ggg tat aca aca gtg acc gga ttg cgg ttg cgt aac	2994
Val Ala Val Asn Gly Tyr Thr Thr Val Thr Gly Leu Arg Leu Arg Asn	
385 390 395 400	
acc aca acg gga gag gaa acc acg cta gta gtg acc ggg gtt ttt gtt	3042
Thr Thr Thr Gly Glu Glu Thr Thr Leu Val Val Thr Gly Val Phe Val	
405 410 415	
gca att ggc cat gaa cca cgt tcc agc ctg gtg agc gat gtc gtc gac	3090
Ala Ile Gly His Glu Pro Arg Ser Ser Leu Val Ser Asp Val Val Asp	
420 425 430	
ata gac ccg gat ggc tac gtc ctg gtg aaa gga cgt acg acg agt aca	3138
Ile Asp Pro Asp Gly Tyr Val Leu Val Lys Gly Arg Thr Thr Ser Thr	
435 440 445	
tcg atg gac ggc gtt ttt gcg gcc ggc gac ctg gta gat cgc acc tac	3186
Ser Met Asp Gly Val Phe Ala Ala Gly Asp Leu Val Asp Arg Thr Tyr	
450 455 460	
cgg cag gcg atc act gcc gca ggt agt gcc tgt gcc gcc gcc atc gac	3234
Arg Gln Ala Ile Thr Ala Ala Gly Ser Gly Cys Ala Ala Ala Ile Asp	

465	470	475	480	
gcc gaa cgt tgg ttg gcg gag cat gcc ggg tca aaa gct aac gaa aca				3282
Ala Glu Arg Trp Leu Ala Glu His Ala Gly Ser Lys Ala Asn Glu Thr	485	490	495	
aca gag gaa act gga gac gtt gac agt acc gac aca acc gat tgg agc				3330
Thr Glu Glu Thr Gly Asp Val Asp Ser Thr Asp Thr Thr Asp Trp Ser	500	505	510	
act gcg atg act gac gcc aag aac gcc ggg gtc aca ata gaa gtg acc				3378
Thr Ala Met Thr Asp Ala Lys Asn Ala Gly Val Thr Ile Glu Val Thr	515	520	525	
gat gct tcc ttt ttc gca gac gtc tta tcc agt aat aag cct gtg tta				3426
Asp Ala Ser Phe Phe Ala Asp Val Leu Ser Ser Asn Lys Pro Val Leu	530	535	540	
gtt gat ttt tgg gca aca tgg tgt gga ccc tgc aag atg gta gcg ccg				3474
Val Asp Phe Trp Ala Thr Trp Cys Gly Pro Cys Lys Met Val Ala Pro	545	550	555	560
gta ctc gaa gag atc gcg tcc gaa caa cga aac cag ctc act gtc gcc				3522
Val Leu Glu Glu Ile Ala Ser Glu Gln Arg Asn Gln Leu Thr Val Ala	565	570	575	
aag tta gat gta gac acc aac ccg gaa atg gca cgc gag ttc cag gtc				3570
Lys Leu Asp Val Asp Thr Asn Pro Glu Met Ala Arg Glu Phe Gln Val	580	585	590	
gtg tcg ata ccc aca atg att ctg ttc cag ggt ggc caa cca gta aaa				3618
Val Ser Ile Pro Thr Met Ile Leu Phe Gln Gly Gly Gln Pro Val Lys	595	600	605	
cgc atc gtt ggc gct aag ggc aaa gca gcg tta cta cgt gac ctt tcc				3666
Arg Ile Val Gly Ala Lys Gly Lys Ala Ala Leu Leu Arg Asp Leu Ser	610	615	620	
gac gtg gta cct aac ctc aat taa gctttaaata agtatgaact aaaatgcatg				3720
Asp Val Val Pro Asn Leu Asn *	625	630		
taggtgtaag agctcatgga gagcatggaa tattgtatcc gaccatgtaa cagtataata				3780
actgagctcc atctcacttc ttctatgaat aaacaaagga tggtatgata tattaacact				3840
ctatctatgc accttattgt tctatgataa atttcctctt attattataa atcatctgaa				3900
tcgtgacggc ttatggaatg cttcaaatag tacaaaaaca aatgtgtact ataagacttt				3960
ctaaacaatt ctaacttttag cattgtgaac gagacataag tgtaagaag acataacaat				4020
tataatggaa gaagtttgtc tccatttata tattatatat taccactta tgtattatat				4080
taggatgta aggagacata acaattataa agagagaagt ttgtatccat ttatatatta				4140
tatactaccc atttatatat tatacttatc cacttattta atgtctttat aaggtttgat				4200
ccatgatatt tctaataatt tagttgatat gtatatgaaa gggactatt tgaactctct				4260
tactctgtat aaaggttgga tcatccttaa agtgggtcta ttaatttta ttgcttctta				4320
cagataaaaa aaaaattatg agttggtttg ataaaaatatt gaaggattta aaataataat				4380
aaataataaa taacatataa tatatgtata taaattttatt ataataaac atttatctat				4440
aaaaaagtaa atattgtcat aaatctatac aatcgtttag ccttgctgga cgactctcaa				4500
ttattttaaac gagagtaaac atatttgact ttttggttat ttaacaaatt attatttaac				4560
actatatgaa attttttttt tttatcgcca aggaaataaa attaaattag gagggacaat				4620
ggtgtgtccc aatccttata caaccaactt ccacaggaag gtcaggtcgg ggacaacaaa				4680
aaaacaggca agggaaattt tttaatttgg gttgtcttgt ttgctgcata atttatgcag				4740
taaaacacta cacataaccc ttttagcagt agagcaatgg ttgaccgtgt gcttagcttc				4800
ttttatttta tttttttatc agcaaagaat aaataaaaata aaatgagaca cttcagggat				4860
gtttcaaccc ttatacaaaa ccccaaaaac aagtttccta gcaccctacc aactaaggta				4920
cc				4922

<210> 31
 <211> 118
 <212> PRT

<213> Artificial Sequence

<220>

<223> Chimeric

<400> 31

```

Met Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp
 1           5           10           15
Gln Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly
      20           25           30
Arg Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr
      35           40           45
Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu
      50           55           60
Val Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile
      65           70           75           80
Phe Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile
      85           90           95
Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val
      100          105          110
Phe Ser Trp Ile Tyr Lys
      115

```

<210> 32

<211> 513

<212> PRT

<213> Artificial Sequence

<220>

<223> Chimeric

<400> 32

```

Tyr Ala Thr Gly Glu His Pro Gln Gly Ser Asp Lys Leu Asp Ser Ala
 1           5           10           15
Arg Met Lys Leu Gly Ser Lys Ala Gln Asp Leu Lys Asp Arg Ala Gln
      20           25           30
Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu His Asp Arg Asp Arg Thr
      35           40           45
Arg Gly Gly Gln His Thr Thr Met Asn Thr Thr Pro Ser Ala His Glu
      50           55           60
Thr Ile His Glu Val Ile Val Ile Gly Ser Gly Pro Ala Gly Tyr Thr
      65           70           75           80
Ala Ala Leu Tyr Ala Ala Arg Ala Gln Leu Thr Pro Leu Val Phe Glu
      85           90           95
Gly Thr Ser Phe Gly Gly Ala Leu Met Thr Thr Thr Glu Val Glu Asn
      100          105          110
Tyr Pro Gly Phe Arg Asn Gly Ile Thr Gly Pro Glu Leu Met Asp Asp
      115          120          125
Met Arg Glu Gln Ala Leu Arg Phe Gly Ala Glu Leu Arg Thr Glu Asp
      130          135          140
Val Glu Ser Val Ser Leu Arg Gly Pro Ile Lys Ser Val Val Thr Ala
      145          150          155          160
Glu Gly Gln Thr Tyr Gln Ala Arg Ala Val Ile Leu Ala Met Gly Thr
      165          170          175
Ser Val Arg Tyr Leu Gln Ile Pro Gly Glu Gln Glu Leu Leu Gly Arg
      180          185          190
Gly Val Ser Ala Cys Ala Thr Cys Asp Gly Ser Phe Phe Arg Gly Gln
      195          200          205
Asp Ile Ala Val Ile Gly Gly Asp Ser Ala Met Glu Glu Ala Leu
      210          215          220
Phe Leu Thr Arg Phe Ala Arg Ser Val Thr Leu Val His Arg Arg Asp
      225          230          235          240
Glu Phe Arg Ala Ser Lys Ile Met Leu Gly Arg Ala Arg Asn Asn Asp
      245          250          255
Lys Ile Lys Phe Ile Thr Asn His Thr Val Val Ala Val Asn Gly Tyr
      260          265          270

```

Thr Thr Val Thr Gly Leu Arg Leu Arg Asn Thr Thr Thr Gly Glu Glu
 275 280 285
 Thr Thr Leu Val Val Thr Gly Val Phe Val Ala Ile Gly His Glu Pro
 290 295 300
 Arg Ser Ser Leu Val Ser Asp Val Val Asp Ile Asp Pro Asp Gly Tyr
 305 310 315 320
 Val Leu Val Lys Gly Arg Thr Thr Ser Thr Ser Met Asp Gly Val Phe
 325 330 335
 Ala Ala Gly Asp Leu Val Asp Arg Thr Tyr Arg Gln Ala Ile Thr Ala
 340 345 350
 Ala Gly Ser Gly Cys Ala Ala Ala Ile Asp Ala Glu Arg Trp Leu Ala
 355 360 365
 Glu His Ala Gly Ser Lys Ala Asn Glu Thr Thr Glu Glu Thr Gly Asp
 370 375 380
 Val Asp Ser Thr Asp Thr Thr Asp Trp Ser Thr Ala Met Thr Asp Ala
 385 390 395 400
 Lys Asn Ala Gly Val Thr Ile Glu Val Thr Asp Ala Ser Phe Phe Ala
 405 410 415
 Asp Val Leu Ser Ser Asn Lys Pro Val Leu Val Asp Phe Trp Ala Thr
 420 425 430
 Trp Cys Gly Pro Cys Lys Met Val Ala Pro Val Leu Glu Glu Ile Ala
 435 440 445
 Ser Glu Gln Arg Asn Gln Leu Thr Val Ala Lys Leu Asp Val Asp Thr
 450 455 460
 Asn Pro Glu Met Ala Arg Glu Phe Gln Val Val Ser Ile Pro Thr Met
 465 470 475 480
 Ile Leu Phe Gln Gly Gly Gln Pro Val Lys Arg Ile Val Gly Ala Lys
 485 490 495
 Gly Lys Ala Ala Leu Leu Arg Asp Leu Ser Asp Val Val Pro Asn Leu
 500 505 510
 Asn

<210> 33
 <211> 4935
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> CDS
 <222> (1554) ... (1906)

<221> CDS
 <222> (2147) ... (3701)

<223> Chimeric

<400> 33
 ctgcaggaat tcattgtact cccagtatca ttatagtgaag agttttggct ctctcgccgg 60
 tggtttttta cctctattta aaggggtttt ccacctaaaa attctggtat cattctcact 120
 ttacttggtta ctttaatttc tcataatctt tggttgaaat tatcacgctt ccgcacacga 180
 tatccctaca aattttattat ttgttaaaca ttttcaaacc gcataaaatt ttatgaagtc 240
 ccgtctatct ttaatgtagt ctaacatttt catattgaaa tatataattt acttaatttt 300
 agcgttggtta gaaagcataa agatttatct ttattcttct tcatataaat gtttaatatata 360
 caatataaac aaattcttta ccttaagaag gatttcccat tttatatattt aaaaatatata 420
 ttatcaaata tttttcaacc acgtaaatct cataataata agttgtttca aaagtaataa 480
 aatttaactc cataattttt ttattcgact gatcttaaag caacaccag tgacacaact 540
 agccattttt ttctttgaat aaaaaaatcc aattatcatt gtattttttt tatacaatga 600
 aaatttcacc aaacaatcat ttgtggtatt tctgaagcaa gtcattgtat gcaaaattct 660
 ataattccca tttgacacta cggaagtaac tgaagatctg cttttacatg cgagacacat 720
 cttctaaagt aattttaata atagttacta tattcaagat ttcataatc aaataactca 780
 tattacttct aaaaaattaa ttagatataa ttaaaatatt acttttttaa ttttaagttt 840
 aattgttgaa tttgtgacta ttgatttatt attctactat gtttaaatgt ttttatagat 900
 agtttaaagt aaatataagt aatgtagtag agtggttagag tgttacccta aaccataaac 960
 tataacattt atggtggact aattttcata tatttcttat tgcttttacc ttttcttggg 1020
 atgtaagtcc gtaactagaa ttacagtggg ttgccatggc actctgtggt ctttttggtt 1080

```

atgcatgggt cttgcgcaag aaaaagacaa agaacaaaga aaaaagacaa aacagagaga 1140
caaaacgcaa tcacacaacc aactcaaatt agtcactggc tgatcaagat cgccgcgtcc 1200
atgtatgtct aaatgccatg caaagcaaca cgtgcttaac atgcacttta aatggctcac 1260
ccatctcaac ccacacacaa acacattgcc tttttcttca tcatcaccac aaccacctgt 1320
atatattcat tctcttcgcg cactcaatt tcttcacttc aacacacgct aacctgcata 1380
tgcggtgcat cccatgccc aatctccatg catgttccaa ccaccttctc tcttatataa 1440
tacctataaa tacctcta atcactcact tctttcatca tccatccatc cagagtacta 1500
ctactctact actataatac cccaacccaa ctcatattca atactactct act atg 1556
Met
1

gcg gat aca gct aga gga acc cat cac gat atc atc ggc aga gac cag 1604
Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp Gln
5 10 15

tac ccg atg atg ggc cga gac cga gac cag tac cag atg tcc gga cga 1652
Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly Arg
20 25 30

gga tct gac tac tcc aag tct agg cag att gct aaa gct gca act gct 1700
Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr Ala
35 40 45

gtc aca gct ggt ggt tcc ctc ctt gtt ctc tcc agc ctt acc ctt gtt 1748
Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu Val
50 55 60 65

gga act gtc ata gct ttg act gtt gca aca cct ctg ctc gtt atc ttc 1796
Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile Phe
70 75 80

agc cca atc ctt gtc ccg gct ctc atc aca gtt gca ctc ctc atc acc 1844
Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile Thr
85 90 95

ggg ttt ctt tcc tct gga ggg ttt ggc att gcc gct ata acc gtt ttc 1892
Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val Phe
100 105 110

tct tgg att tac aa gtaagcacac atttatcatc ttacttcata attttgtgca 1946
Ser Trp Ile Tyr Lys
115

atatgtgcat gcatgtgttg agccagtagc tttggatcaa tttttttggt cgaataacaa 2006
atgtaacaat aagaaattgc aaattctagg gaacatttgg ttaactaaat acgaaatttg 2066
acctagctag cttgaatgtg tctgtgtata tcatctatat aggtaaaatg cttgggtatga 2126
tacctattga ttgtgaatag g tac gca acg gga gag cac cca cag gga tca 2177
Tyr Ala Thr Gly Glu His Pro Gln Gly Ser
120 125

gac aag ttg gac agt gca agg atg aag ttg gga agc aaa gct cag gat 2225
Asp Lys Leu Asp Ser Ala Arg Met Lys Leu Gly Ser Lys Ala Gln Asp
130 135 140

ctg aaa gac aga gct cag tac tac gga cag caa cat act ggt ggg gaa 2273
Leu Lys Asp Arg Ala Gln Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu
145 150 155 160

cat gac cgt gac cgt act cgt ggt ggc cag cac act acc atg aat ggt 2321
His Asp Arg Asp Arg Thr Arg Gly Gly Gln His Thr Thr Met Asn Gly
165 170 175

ctc gaa act cac aac aca agg ctc tgt atc gta gga agt ggc cca gcg 2369
Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser Gly Pro Ala
180 185 190

gca cac acg gcg gcg att tac gca gct agg gct gaa ctt aaa cct ctt 2417

```

Ala	His	Thr	Ala	Ala	Ile	Tyr	Ala	Ala	Arg	Ala	Glu	Leu	Lys	Pro	Leu		
		195					200					205					
ctc	ttc	gaa	gga	tgg	atg	gct	aac	gac	atc	gct	ccc	ggg	ggg	caa	cta	2465	
Leu	Phe	Glu	Gly	Trp	Met	Ala	Asn	Asp	Ile	Ala	Pro	Gly	Gly	Gln	Leu		
	210					215					220						
aca	acc	acc	acc	gac	gtc	gag	aac	ttc	ccc	gga	ttt	cca	gaa	ggg	att	2513	
Thr	Thr	Thr	Thr	Asp	Val	Glu	Asn	Phe	Pro	Gly	Phe	Pro	Glu	Gly	Ile		
225					230					235					240		
ctc	gga	gta	gag	ctc	act	gac	aaa	ttc	cgt	aaa	caa	tcg	gag	cga	ttc	2561	
Leu	Gly	Val	Glu	Leu	Thr	Asp	Lys	Phe	Arg	Lys	Gln	Ser	Glu	Arg	Phe		
				245					250					255			
ggg	act	acg	ata	ttt	aca	gag	acg	gtg	acg	aaa	gtc	gat	ttc	tct	tcg	2609	
Gly	Thr	Thr	Ile	Phe	Thr	Glu	Thr	Val	Thr	Lys	Val	Asp	Phe	Ser	Ser		
			260					265					270				
aaa	ccg	ttt	aag	cta	ttc	aca	gat	tca	aaa	gcc	att	ctc	gct	gac	gct	2657	
Lys	Pro	Phe	Lys	Leu	Phe	Thr	Asp	Ser	Lys	Ala	Ile	Leu	Ala	Asp	Ala		
	275						280					285					
gtg	att	ctc	gct	act	gga	gct	gtg	gct	aag	cgg	ctt	agc	ttc	gtt	gga	2705	
Val	Ile	Leu	Ala	Thr	Gly	Ala	Val	Ala	Lys	Arg	Leu	Ser	Phe	Val	Gly		
	290					295					300						
tct	ggg	gaa	ggg	tct	gga	ggg	ttc	tgg	aac	cgt	gga	atc	tcc	gct	tgt	2753	
Ser	Gly	Glu	Gly	Ser	Gly	Gly	Phe	Trp	Asn	Arg	Gly	Ile	Ser	Ala	Cys		
305					310				315						320		
gct	gtt	tgc	gac	gga	gct	gct	ccg	ata	ttc	cgt	aac	aaa	cct	ctt	gcg	2801	
Ala	Val	Cys	Asp	Gly	Ala	Ala	Pro	Ile	Phe	Arg	Asn	Lys	Pro	Leu	Ala		
				325					330					335			
gtg	atc	ggg	gga	ggc	gat	tca	gca	atg	gaa	gaa	gca	aac	ttt	ctt	aca	2849	
Val	Ile	Gly	Gly	Gly	Asp	Ser	Ala	Met	Glu	Glu	Ala	Asn	Phe	Leu	Thr		
			340					345					350				
aaa	tat	gga	tct	aaa	gtg	tat	ata	atc	cat	agg	aga	gat	gct	ttt	aga	2897	
Lys	Tyr	Gly	Ser	Lys	Val	Tyr	Ile	Ile	His	Arg	Arg	Asp	Ala	Phe	Arg		
	355						360					365					
gcg	tct	aag	att	atg	cag	cag	cga	gct	ttg	tct	aat	cct	aag	att	gat	2945	
Ala	Ser	Lys	Ile	Met	Gln	Gln	Arg	Ala	Leu	Ser	Asn	Pro	Lys	Ile	Asp		
	370					375					380						
gtg	att	tgg	aac	tcg	tct	gtt	gtg	gaa	gct	tat	gga	gat	gga	gaa	aga	2993	
Val	Ile	Trp	Asn	Ser	Ser	Val	Val	Glu	Ala	Tyr	Gly	Asp	Gly	Glu	Arg		
385					390				395						400		
gat	gtg	ctt	gga	gga	ttg	aaa	gtg	aag	aat	gtg	gtt	acc	gga	gat	gtt	3041	
Asp	Val	Leu	Gly	Gly	Leu	Lys	Val	Lys	Asn	Val	Val	Thr	Gly	Asp	Val		
			405					410					415				
tct	gat	tta	aaa	gtt	tct	gga	ttg	ttc	ttt	gct	att	ggg	cat	gag	cca	3089	
Ser	Asp	Leu	Lys	Val	Ser	Gly	Leu	Phe	Phe	Ala	Ile	Gly	His	Glu	Pro		
			420					425					430				
gct	acc	aag	ttt	ttg	gat	ggg	ggg	gtt	gag	tta	gat	tcg	gat	ggg	tat	3137	
Ala	Thr	Lys	Phe	Leu	Asp	Gly	Gly	Val	Glu	Leu	Asp	Ser	Asp	Gly	Tyr		
	435					440					445						
gtt	gtc	acg	aag	cct	ggg	act	aca	cag	act	agc	gtt	ccc	gga	gtt	ttc	3185	
Val	Val	Thr	Lys	Pro	Gly	Thr	Thr	Gln	Thr	Ser	Val	Pro	Gly	Val	Phe		
	450					455					460						

gct gcg ggt gat gtt cag gat aag aag tat agg caa gcc atc act gct	3233
Ala Ala Gly Asp Val Gln Asp Lys Lys Tyr Arg Gln Ala Ile Thr Ala	
465 470 475 480	
gca gga act ggg tgc atg gca gct ttg gat gca gag cat tac tta caa	3281
Ala Gly Thr Gly Cys Met Ala Ala Leu Asp Ala Glu His Tyr Leu Gln	
485 490 495	
gag att gct gga tcg aag gct aac gag acc acc gag gaa act gga gat	3329
Glu Ile Ala Gly Ser Lys Ala Asn Glu Thr Thr Glu Glu Thr Gly Asp	
500 505 510	
gtt gac tcg acg gat act acg gat tgg tcg acg gct atg gaa gaa gga	3377
Val Asp Ser Thr Asp Thr Thr Asp Trp Ser Thr Ala Met Glu Glu Gly	
515 520 525	
caa gtg atc gcc tgc cac acc gtt gag aca tgg aac gag cag ctt cag	3425
Gln Val Ile Ala Cys His Thr Val Glu Thr Trp Asn Glu Gln Leu Gln	
530 535 540	
aag gct aat gaa tcc aaa act ctt gtg gtg gtt gat ttc acg gct tct	3473
Lys Ala Asn Glu Ser Lys Thr Leu Val Val Val Asp Phe Thr Ala Ser	
545 550 555 560	
tgg tgt gga cca tgt cgt ttc atc gct cca ttc ttt gct gat ttg gct	3521
Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro Phe Phe Ala Asp Leu Ala	
565 570 575	
aag aaa ctt cct aac gtg ctt ttc ctc aag gtt gat act gat gaa ttg	3569
Lys Lys Leu Pro Asn Val Leu Phe Leu Lys Val Asp Thr Asp Glu Leu	
580 585 590	
aag tcg gtg gca agt gat tgg gcg ata cag gcg atg cca acc ttc atg	3617
Lys Ser Val Ala Ser Asp Trp Ala Ile Gln Ala Met Pro Thr Phe Met	
595 600 605	
ttt ttg aag gaa ggg aag att ttg gac aaa gtt gtt gga gcc aag aaa	3665
Phe Leu Lys Glu Gly Lys Ile Leu Asp Lys Val Val Gly Ala Lys Lys	
610 615 620	
gat gag ctt cag tct acc att gcc aaa cac ttg gct taagcttaaa	3711
Asp Glu Leu Gln Ser Thr Ile Ala Lys His Leu Ala	
625 630 635	
taagtatgaa ctaaaatgca tgtaggtgta agagctcatg gagagcatgg aatattgtat	3771
ccgaccatgt aacagtataa taactgagct ccattctact tcttctatga ataaacaaag	3831
gatgttatga tatattaaca ctctatctat gcaccttatt gttctatgat aaatttcctc	3891
ttattattat aaatcatctg aatcgtgacg gcttatggaa tgcttcaaat agtacaaaaa	3951
caaatgtgta ctataagact ttctaaacaa ttctaacttt agcattgtga acgagacata	4011
agtgttaaga agacataaca attataatgg aagaagtttg tctccattta tatattatat	4071
attaccctact tatgtattat attaggatgt taaggagaca taacaattat aaagagagaa	4131
gtttgtatcc atttatatat tatatactac ccatttatat attatactta tccacttatt	4191
taatgtcttt ataaggtttg atccatgata tttctaatat tttagttgat atgtatatga	4251
aagggtacta tttgaactct cttactctgt ataaaggttg gatcatcctt aaagtgggtc	4311
tatttaatttt tattgcttct tacagataaa aaaaaaatta tgagttgggt tgataaaaata	4371
ttgaaggatt taaaataata ataaataata aataacatat aatatatgta tataaattta	4431
ttataatata acatttatct ataaaaaagt aaatatgtc ataatcttat acaatcgttt	4491
agccttgctg gacgactctc aattatttaa acgagagtaa acatatttga ctttttggtt	4551
atttaacaaa ttattattta acactatatg aaattttttt tttttatcgg caaggaaata	4611
aaattaaatt aggagggaca atgggtgtgtc ccaatcotta tacaaccaac ttccacagga	4671
aggtcagggtc ggggacaaca aaaaaacagg caagggaaat tttttaattt gggttgctt	4731
gtttgctgca taatttatgc agtaaacac tacacataac ctttttagca gtagagcaat	4791
ggttgaccgt gtgcttagct tcttttattt tattttttta tcagcaaaga ataaataaaa	4851
taaaatgaga cacttcaggg atgtttcaac ccttatataa aaccccaaaa acaagtttcc	4911
tagcacccta ccaactaagg tacc	4935

<210> 34

<211> 118
 <212> PRT
 <213> Artificial Sequence

<220>
 <221> SITE
 <222> (1)...(118)
 <223> oleosin

<223> Chimeric

<400> 34
 Met Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp
 1 5 10 15
 Gln Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly
 20 25 30
 Arg Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr
 35 40 45
 Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu
 50 55 60
 Val Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile
 65 70 75 80
 Phe Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile
 85 90 95
 Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val
 100 105 110
 Phe Ser Trp Ile Tyr Lys
 115

<210> 35
 <211> 518
 <212> PRT
 <213> Artificial Sequence

<220>
 <221> SITE
 <222> (1)...(55)
 <223> oleosin

<221> SITE
 <222> (56)...(383)
 <223> thioredoxin reductase

<221> SITE
 <222> (384)...(406)
 <223> linker

<221> SITE
 <222> (407)...(518)
 <223> thioredoxin

<223> Chimeric

<400> 35
 Tyr Ala Thr Gly Glu His Pro Gln Gly Ser Asp Lys Leu Asp Ser Ala
 1 5 10 15
 Arg Met Lys Leu Gly Ser Lys Ala Gln Asp Leu Lys Asp Arg Ala Gln
 20 25 30
 Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu His Asp Arg Asp Arg Thr
 35 40 45
 Arg Gly Gly Gln His Thr Thr Met Asn Gly Leu Glu Thr His Asn Thr
 50 55 60
 Arg Leu Cys Ile Val Gly Ser Gly Pro Ala Ala His Thr Ala Ala Ile
 65 70 75 80
 Tyr Ala Ala Arg Ala Glu Leu Lys Pro Leu Leu Phe Glu Gly Trp Met
 85 90 95

Ala Asn Asp Ile Ala Pro Gly Gly Gln Leu Thr Thr Thr Thr Asp Val
 100 105 110
 Glu Asn Phe Pro Gly Phe Pro Glu Gly Ile Leu Gly Val Glu Leu Thr
 115 120 125
 Asp Lys Phe Arg Lys Gln Ser Glu Arg Phe Gly Thr Thr Ile Phe Thr
 130 135 140
 Glu Thr Val Thr Lys Val Asp Phe Ser Ser Lys Pro Phe Lys Leu Phe
 145 150 155 160
 Thr Asp Ser Lys Ala Ile Leu Ala Asp Ala Val Ile Leu Ala Thr Gly
 165 170 175
 Ala Val Ala Lys Arg Leu Ser Phe Val Gly Ser Gly Glu Gly Ser Gly
 180 185 190
 Gly Phe Trp Asn Arg Gly Ile Ser Ala Cys Ala Val Cys Asp Gly Ala
 195 200 205
 Ala Pro Ile Phe Arg Asn Lys Pro Leu Ala Val Ile Gly Gly Gly Asp
 210 215 220
 Ser Ala Met Glu Glu Ala Asn Phe Leu Thr Lys Tyr Gly Ser Lys Val
 225 230 235 240
 Tyr Ile Ile His Arg Arg Asp Ala Phe Arg Ala Ser Lys Ile Met Gln
 245 250 255
 Gln Arg Ala Leu Ser Asn Pro Lys Ile Asp Val Ile Trp Asn Ser Ser
 260 265 270
 Val Val Glu Ala Tyr Gly Asp Gly Glu Arg Asp Val Leu Gly Gly Leu
 275 280 285
 Lys Val Lys Asn Val Val Thr Gly Asp Val Ser Asp Leu Lys Val Ser
 290 295 300
 Gly Leu Phe Phe Ala Ile Gly His Glu Pro Ala Thr Lys Phe Leu Asp
 305 310 315 320
 Gly Gly Val Glu Leu Asp Ser Asp Gly Tyr Val Val Thr Lys Pro Gly
 325 330 335
 Thr Thr Gln Thr Ser Val Pro Gly Val Phe Ala Ala Gly Asp Val Gln
 340 345 350
 Asp Lys Lys Tyr Arg Gln Ala Ile Thr Ala Ala Gly Thr Gly Cys Met
 355 360 365
 Ala Ala Leu Asp Ala Glu His Tyr Leu Gln Glu Ile Ala Gly Ser Lys
 370 375 380
 Ala Asn Glu Thr Thr Glu Glu Thr Gly Asp Val Asp Ser Thr Asp Thr
 385 390 395 400
 Thr Asp Trp Ser Thr Ala Met Glu Glu Gly Gln Val Ile Ala Cys His
 405 410 415
 Thr Val Glu Thr Trp Asn Glu Gln Leu Gln Lys Ala Asn Glu Ser Lys
 420 425 430
 Thr Leu Val Val Val Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg
 435 440 445
 Phe Ile Ala Pro Phe Phe Ala Asp Leu Ala Lys Lys Leu Pro Asn Val
 450 455 460
 Leu Phe Leu Lys Val Asp Thr Asp Glu Leu Lys Ser Val Ala Ser Asp
 465 470 475 480
 Trp Ala Ile Gln Ala Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys
 485 490 495
 Ile Leu Asp Lys Val Val Gly Ala Lys Lys Asp Glu Leu Gln Ser Thr
 500 505 510
 Ile Ala Lys His Leu Ala
 515

<210> 36

<211> 458

<212> PRT

<213> Mycobacterium leprae

<400> 36

Met Asn Thr Thr Pro Ser Ala His Glu Thr Ile His Glu Val Ile Val
 1 5 10 15
 Ile Gly Ser Gly Pro Ala Gly Tyr Thr Ala Ala Leu Tyr Ala Ala Arg
 20 25 30
 Ala Gln Leu Thr Pro Leu Val Phe Glu Gly Thr Ser Phe Gly Gly Ala

```

      35      40      45
Leu Met Thr Thr Thr Glu Val Glu Asn Tyr Pro Gly Phe Arg Asn Gly
  50      55      60
Ile Thr Gly Pro Glu Leu Met Asp Asp Met Arg Glu Gln Ala Leu Arg
  65      70      75
Phe Gly Ala Glu Leu Arg Thr Glu Asp Val Glu Ser Val Ser Leu Arg
      85      90      95
Gly Pro Ile Lys Ser Val Val Thr Ala Glu Gly Gln Thr Tyr Gln Ala
      100      105      110
Arg Ala Val Ile Leu Ala Met Gly Thr Ser Val Arg Tyr Leu Gln Ile
      115      120      125
Pro Gly Glu Gln Glu Leu Leu Gly Arg Gly Val Ser Ala Cys Ala Thr
      130      135      140
Cys Asp Gly Ser Phe Phe Arg Gly Gln Asp Ile Ala Val Ile Gly Gly
      145      150      155
Gly Asp Ser Ala Met Glu Glu Ala Leu Phe Leu Thr Arg Phe Ala Arg
      165      170      175
Ser Val Thr Leu Val His Arg Arg Asp Glu Phe Arg Ala Ser Lys Ile
      180      185      190
Met Leu Gly Arg Ala Arg Asn Asn Asp Lys Ile Lys Phe Ile Thr Asn
      195      200      205
His Thr Val Val Ala Val Asn Gly Tyr Thr Thr Val Thr Gly Leu Arg
      210      215      220
Leu Arg Asn Thr Thr Thr Gly Glu Glu Thr Thr Leu Val Val Thr Gly
      225      230      235
Val Phe Val Ala Ile Gly His Glu Pro Arg Ser Ser Leu Val Ser Asp
      245      250      255
Val Val Asp Ile Asp Pro Asp Gly Tyr Val Leu Val Lys Gly Arg Thr
      260      265      270
Thr Ser Thr Ser Met Asp Gly Val Phe Ala Ala Gly Asp Leu Val Asp
      275      280      285
Arg Thr Tyr Arg Gln Ala Ile Thr Ala Ala Gly Ser Gly Cys Ala Ala
      290      295      300
Ala Ile Asp Ala Glu Arg Trp Leu Ala Glu His Ala Gly Ser Lys Ala
      305      310      315
Asn Glu Thr Thr Glu Glu Thr Gly Asp Val Asp Ser Thr Asp Thr Thr
      325      330      335
Asp Trp Ser Thr Ala Met Thr Asp Ala Lys Asn Ala Gly Val Thr Ile
      340      345      350
Glu Val Thr Asp Ala Ser Phe Phe Ala Asp Val Leu Ser Ser Asn Lys
      355      360      365
Pro Val Leu Val Asp Phe Trp Ala Thr Trp Cys Gly Pro Cys Lys Met
      370      375      380
Val Ala Pro Val Leu Glu Glu Ile Ala Ser Glu Gln Arg Asn Gln Leu
      385      390      395
Thr Val Ala Lys Leu Asp Val Asp Thr Asn Pro Glu Met Ala Arg Glu
      405      410      415
Phe Gln Val Val Ser Ile Pro Thr Met Ile Leu Phe Gln Gly Gly Gln
      420      425      430
Pro Val Lys Arg Ile Val Gly Ala Lys Gly Lys Ala Ala Leu Leu Arg
      435      440      445
Asp Leu Ser Asp Val Val Pro Asn Leu Asn
      450      455

```

<210> 37

<211> 471

<212> PRT

<213> Arabidopsis thaliana

<220>

<223> Chimeric

<400> 37

```

Met Asn Gly Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser
  1      5      10      15
Gly Pro Ala Ala His Thr Ala Ala Ile Tyr Ala Ala Arg Ala Glu Leu

```


			20					25				30			
Lys	Pro	Leu	Leu	Phe	Glu	Gly	Trp	Met	Ala	Asn	Asp	Ile	Ala	Pro	Gly
		35					40					45			
Gly	Gln	Leu	Thr	Thr	Thr	Thr	Asp	Val	Glu	Asn	Phe	Pro	Gly	Phe	Pro
	50					55					60				
Glu	Gly	Ile	Leu	Gly	Val	Glu	Leu	Thr	Asp	Lys	Phe	Arg	Lys	Gln	Ser
65					70					75					80
Glu	Arg	Phe	Gly	Thr	Thr	Ile	Phe	Thr	Glu	Thr	Val	Thr	Lys	Val	Asp
				85					90					95	
Phe	Ser	Ser	Lys	Pro	Phe	Lys	Leu	Phe	Thr	Asp	Ser	Lys	Ala	Ile	Leu
			100					105					110		
Ala	Asp	Ala	Val	Ile	Leu	Ala	Thr	Gly	Ala	Val	Ala	Lys	Arg	Leu	Ser
		115					120					125			
Phe	Val	Gly	Ser	Gly	Glu	Gly	Ser	Gly	Gly	Phe	Trp	Asn	Arg	Gly	Ile
	130					135					140				
Ser	Ala	Cys	Ala	Val	Cys	Asp	Gly	Ala	Ala	Pro	Ile	Phe	Arg	Asn	Lys
145					150					155					160
Pro	Leu	Ala	Val	Ile	Gly	Gly	Gly	Asp	Ser	Ala	Met	Glu	Glu	Ala	Asn
				165					170					175	
Phe	Leu	Thr	Lys	Tyr	Gly	Ser	Lys	Val	Tyr	Ile	Ile	His	Arg	Arg	Asp
			180					185					190		
Ala	Phe	Arg	Ala	Ser	Lys	Ile	Met	Gln	Gln	Arg	Ala	Leu	Ser	Asn	Pro
		195					200					205			
Lys	Ile	Asp	Val	Ile	Trp	Asn	Ser	Ser	Val	Val	Glu	Ala	Tyr	Gly	Asp
	210					215					220				
Gly	Glu	Arg	Asp	Val	Leu	Gly	Gly	Leu	Lys	Val	Lys	Asn	Val	Val	Thr
225					230					235					240
Gly	Asp	Val	Ser	Asp	Leu	Lys	Val	Ser	Gly	Leu	Phe	Phe	Ala	Ile	Gly
				245					250				255		
His	Glu	Pro	Ala	Thr	Lys	Phe	Leu	Asp	Gly	Gly	Val	Glu	Leu	Asp	Ser
			260					265					270		
Asp	Gly	Tyr	Val	Val	Thr	Lys	Pro	Gly	Thr	Thr	Gln	Thr	Ser	Val	Pro
		275					280					285			
Gly	Val	Phe	Ala	Ala	Gly	Asp	Val	Gln	Asp	Lys	Lys	Tyr	Arg	Gln	Ala
	290					295					300				
Ile	Thr	Ala	Ala	Gly	Thr	Gly	Cys	Met	Ala	Ala	Leu	Asp	Ala	Glu	His
305					310					315					320
Tyr	Leu	Gln	Glu	Ile	Ala	Gly	Ser	Lys	Ala	Asn	Glu	Thr	Thr	Glu	Glu
				325					330					335	
Thr	Gly	Asp	Val	Asp	Ser	Thr	Asp	Thr	Thr	Asp	Trp	Ser	Thr	Ala	Met
			340					345					350		
Glu	Glu	Gly	Gln	Val	Ile	Ala	Cys	Glu	Glu	Gly	Gln	Val	Ile	Ala	Cys
		355					360								

```
<210> 38
<211> 345
<212> DNA
<213> Arabidopsis thaliana
```

```
<220>
<221> CDS
<222> (1) ... (345)
```

<400> 38
 atg gct tcg gaa gaa gga caa gtg atc gcc tgc cac acc gtt gag aca 48
 Met Ala Ser Glu Glu Gly Gln Val Ile Ala Cys His Thr Val Glu Thr
 1 5 10 15
 tgg aac gag cag ctt cag aag gct aat gaa tcc aaa act ctt gtg gtg 96
 Trp Asn Glu Gln Leu Gln Lys Ala Asn Glu Ser Lys Thr Leu Val Val
 20 25 30
 gtt gat ttc acg gct tct tgg tgt gga cca tgt cgt ttc atc gct cca 144
 Val Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro
 35 40 45
 ttc ttt gct gat ttg gct aag aaa ctt cct aac gtg ctt ttc ctc aag 192
 Phe Phe Ala Asp Leu Ala Lys Lys Leu Pro Asn Val Leu Phe Leu Lys
 50 55 60
 gtt gat act gat gaa ttg aag tcg gtg gca agt gat tgg gcg ata cag 240
 Val Asp Thr Asp Glu Leu Lys Ser Val Ala Ser Asp Trp Ala Ile Gln
 65 70 75 80
 gcg atg cca acc ttc atg ttt ttg aag gaa ggg aag att ttg gac aaa 288
 Ala Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Leu Asp Lys
 85 90 95
 gtt gtt gga gcc aag aaa gat gag ctt cag tct acc att gcc aaa cac 336
 Val Val Gly Ala Lys Lys Asp Glu Leu Gln Ser Thr Ile Ala Lys His
 100 105 110
 ttg gct taa 345
 Leu Ala *

<210> 39
 <211> 114
 <212> PRT
 <213> Arabidopsis thaliana

<400> 39
 Met Ala Ser Glu Glu Gly Gln Val Ile Ala Cys His Thr Val Glu Thr
 1 5 10 15
 Trp Asn Glu Gln Leu Gln Lys Ala Asn Glu Ser Lys Thr Leu Val Val
 20 25 30
 Val Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro
 35 40 45
 Phe Phe Ala Asp Leu Ala Lys Lys Leu Pro Asn Val Leu Phe Leu Lys
 50 55 60
 Val Asp Thr Asp Glu Leu Lys Ser Val Ala Ser Asp Trp Ala Ile Gln
 65 70 75 80
 Ala Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Leu Asp Lys
 85 90 95
 Val Val Gly Ala Lys Lys Asp Glu Leu Gln Ser Thr Ile Ala Lys His
 100 105 110
 Leu Ala

<210> 40
 <211> 999
 <212> DNA
 <213> Arabidopsis thaliana

<220>
 <221> CDS
 <222> (1) ... (999)

<400> 40
 atg aat ggt ctc gaa act cac aac aca agg ctc tgt atc gta gga agt 48
 Met Asn Gly Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser
 1 5 10 15

ggc cca gcg gca cac acg gcg gcg att tac gca gct agg gct gaa ctt 96
 Gly Pro Ala Ala His Thr Ala Ala Ile Tyr Ala Ala Arg Ala Glu Leu
 20 25 30

aaa cct ctt ctc ttc gaa gga tgg atg gct aac gac atc gct ccc ggt 144
 Lys Pro Leu Leu Phe Glu Gly Trp Met Ala Asn Asp Ile Ala Pro Gly
 35 40 45

ggt caa ctc aac caa cca ccg cgt gag aat ttc ccc gga ttt cca gaa 192
 Gly Gln Leu Asn Gln Pro Pro Arg Glu Asn Phe Pro Gly Phe Pro Glu
 50 55 60

ggt att ctc gga gta gag ctc act gac aaa ttc cgt aaa caa tcg gag 240
 Gly Ile Leu Gly Val Glu Leu Thr Asp Lys Phe Arg Lys Gln Ser Glu
 65 70 75 80

cga ttc ggt act acg ata ttt aca gag acg gtg acg aaa gtc gat ttc 288
 Arg Phe Gly Thr Thr Ile Phe Thr Glu Thr Val Thr Lys Val Asp Phe
 85 90 95

tct tcg aaa ccg ttt aag cta ttc aca gat tca aaa gcc att ctc gct 336
 Ser Ser Lys Pro Phe Lys Leu Phe Thr Asp Ser Lys Ala Ile Leu Ala
 100 105 110

gac gct gtg att ctc gct atc gga gct gtg gct aag tgg ctt agc ttc 384
 Asp Ala Val Ile Leu Ala Ile Gly Ala Val Ala Lys Trp Leu Ser Phe
 115 120 125

gtt gga tct ggt gaa gtt ctc gga ggt ttg tgg aac cgt gga atc tcc 432
 Val Gly Ser Gly Glu Val Leu Gly Gly Leu Trp Asn Arg Gly Ile Ser
 130 135 140

gct tgt gct gtt tgc gac gga gct gct ccg ata ttc cgc aac aaa cct 480
 Ala Cys Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg Asn Lys Pro
 145 150 155 160

ctt gcg gtg atc ggt gga ggc gat tct gca atg gaa gaa gca aac ttt 528
 Leu Ala Val Ile Gly Gly Gly Asp Ser Ala Met Glu Glu Ala Asn Phe
 165 170 175

ctt aca aaa tat gga tct aaa gtg tat ata atc gat agg aga gat gct 576
 Leu Thr Lys Tyr Gly Ser Lys Val Tyr Ile Ile Asp Arg Arg Asp Ala
 180 185 190

ttt aga gcg tct aag att atg cag cag cga gct ttg tct aat cct aag 624
 Phe Arg Ala Ser Lys Ile Met Gln Gln Arg Ala Leu Ser Asn Pro Lys
 195 200 205

att gat gtg att tgg aac tcg tct gtt gtg gaa gct tat gga gat gga 672
 Ile Asp Val Ile Trp Asn Ser Ser Val Val Glu Ala Tyr Gly Asp Gly
 210 215 220

gaa aga gat gtg ctt gga gga ttg aaa gtg aag aat gtg gtt acc gga 720
 Glu Arg Asp Val Leu Gly Gly Leu Lys Val Lys Asn Val Val Thr Gly
 225 230 235 240

gat gtt tct gat tta aaa gtt tct gga ttg ttc ttt gct att ggt cat 768
 Asp Val Ser Asp Leu Lys Val Ser Gly Leu Phe Phe Ala Ile Gly His
 245 250 255

gag cca gct acc aag ttt ttg gat ggt ggt gtt gag tta gat tcg gat 816
 Glu Pro Ala Thr Lys Phe Leu Asp Gly Gly Val Glu Leu Asp Ser Asp

260										265					270					
ggt	tat	ggt	gtc	acg	aag	cct	ggt	act	aca	cag	act	agc	ggt	ccc	gga	864				
Gly	Tyr	Val	Val	Thr	Lys	Pro	Gly	Thr	Thr	Gln	Thr	Ser	Val	Pro	Gly					
		275					280					285								
ggt	ttc	gct	gcg	ggg	gat	ggt	cag	gat	aag	aag	tat	agg	caa	gcc	atc	912				
Val	Phe	Ala	Ala	Gly	Asp	Val	Gln	Asp	Lys	Lys	Tyr	Arg	Gln	Ala	Ile					
	290					295					300									
act	gct	gca	gga	act	ggg	tgc	atg	gca	gct	ttg	gat	gca	gag	cat	tac	960				
Thr	Ala	Ala	Gly	Thr	Gly	Cys	Met	Ala	Ala	Leu	Asp	Ala	Glu	His	Tyr					
305					310					315					320					
tta	caa	gag	att	gga	tct	cag	caa	ggt	aag	agt	gat	tga				999				
Leu	Gln	Glu	Ile	Gly	Ser	Gln	Gln	Gly	Lys	Ser	Asp	*								
				325					330											

<210> 41

<211> 332

<212> PRT

<213> Arabidopsis thaliana

<400> 41

Met	Asn	Gly	Leu	Glu	Thr	His	Asn	Thr	Arg	Leu	Cys	Ile	Val	Gly	Ser
1				5					10					15	
Gly	Pro	Ala	Ala	His	Thr	Ala	Ala	Ile	Tyr	Ala	Ala	Arg	Ala	Glu	Leu
		20						25					30		
Lys	Pro	Leu	Phe	Glu	Gly	Trp	Met	Ala	Asn	Asp	Ile	Ala	Pro	Gly	
	35					40					45				
Gly	Gln	Leu	Asn	Gln	Pro	Pro	Arg	Glu	Asn	Phe	Pro	Gly	Phe	Pro	Glu
	50					55					60				
Gly	Ile	Leu	Gly	Val	Glu	Leu	Thr	Asp	Lys	Phe	Arg	Lys	Gln	Ser	Glu
	65				70					75				80	
Arg	Phe	Gly	Thr	Thr	Ile	Phe	Thr	Glu	Thr	Val	Thr	Lys	Val	Asp	Phe
			85					90					95		
Ser	Ser	Lys	Pro	Phe	Lys	Leu	Phe	Thr	Asp	Ser	Lys	Ala	Ile	Leu	Ala
			100					105					110		
Asp	Ala	Val	Ile	Leu	Ala	Ile	Gly	Ala	Val	Ala	Lys	Trp	Leu	Ser	Phe
		115					120					125			
Val	Gly	Ser	Gly	Glu	Val	Leu	Gly	Gly	Leu	Trp	Asn	Arg	Gly	Ile	Ser
	130					135					140				
Ala	Cys	Ala	Val	Cys	Asp	Gly	Ala	Ala	Pro	Ile	Phe	Arg	Asn	Lys	Pro
	145				150					155				160	
Leu	Ala	Val	Ile	Gly	Gly	Gly	Asp	Ser	Ala	Met	Glu	Glu	Ala	Asn	Phe
			165							170				175	
Leu	Thr	Lys	Tyr	Gly	Ser	Lys	Val	Tyr	Ile	Ile	Asp	Arg	Arg	Asp	Ala
			180					185					190		
Phe	Arg	Ala	Ser	Lys	Ile	Met	Gln	Gln	Arg	Ala	Leu	Ser	Asn	Pro	Lys
		195					200					205			
Ile	Asp	Val	Ile	Trp	Asn	Ser	Ser	Val	Val	Glu	Ala	Tyr	Gly	Asp	Gly
	210				215						220				
Glu	Arg	Asp	Val	Leu	Gly	Gly	Leu	Lys	Val	Lys	Asn	Val	Val	Thr	Gly
	225				230					235				240	
Asp	Val	Ser	Asp	Leu	Lys	Val	Ser	Gly	Leu	Phe	Phe	Ala	Ile	Gly	His
				245					250					255	
Glu	Pro	Ala	Thr	Lys	Phe	Leu	Asp	Gly	Gly	Val	Glu	Leu	Asp	Ser	Asp
		260						265					270		
Gly	Tyr	Val	Val	Thr	Lys	Pro	Gly	Thr	Thr	Gln	Thr	Ser	Val	Pro	Gly
		275					280					285			
Val	Phe	Ala	Ala	Gly	Asp	Val	Gln	Asp	Lys	Lys	Tyr	Arg	Gln	Ala	Ile
	290					295					300				
Thr	Ala	Ala	Gly	Thr	Gly	Cys	Met	Ala	Ala	Leu	Asp	Ala	Glu	His	Tyr
	305				310					315				320	
Leu	Gln	Glu	Ile	Gly	Ser	Gln	Gln	Gly	Lys	Ser	Asp				
				325					330						

<210> 42
 <211> 332
 <212> DNA
 <213> E. coli

<220>
 <221> CDS
 <222> (1) ... (332)

```

<400> 42
atg agc gat aaa att att cac ctg act gac gac agt ttt gac acg gat   48
Met Ser Asp Lys Ile Ile His Leu Thr Asp Asp Ser Phe Asp Thr Asp
  1             5             10             15

gta ctc aaa gcg gac ggg gct atc ctc gtt gat ttc tgg gca gag tgg   96
Val Leu Lys Ala Asp Gly Ala Ile Leu Val Asp Phe Trp Ala Glu Trp
             20             25             30

tgc ggg ccg tgt aaa atg atc gct ccg att ctg gat gaa atc gct gac  144
Cys Gly Pro Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp
             35             40             45

gaa tat cag ggc aaa ttg acc gtt gcc aaa ctg aac att gac cag aac  192
Glu Tyr Gln Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn
             50             55             60

cca ggt act gcg cct aaa tat ggc atc cgc ggt att ccg act ctg ctg  240
Pro Gly Thr Ala Pro Lys Tyr Gly Ile Arg Gly Ile Pro Thr Leu Leu
             65             70             75

ctg ttt aaa aac ggc gaa gtg gcg gca acc aaa gta ggc gca ctg tct  288
Leu Phe Lys Asn Gly Glu Val Ala Ala Thr Lys Val Gly Ala Leu Ser
             85             90             95

aaa ggt cag ttg aaa gag ttt ctc gac gcc aat ctg gcg taa ta   332
Lys Gly Gln Leu Lys Glu Phe Leu Asp Ala Asn Leu Ala *
             100             105

```

<210> 43
 <211> 109
 <212> PRT
 <213> E. coli

```

<400> 43
Met Ser Asp Lys Ile Ile His Leu Thr Asp Asp Ser Phe Asp Thr Asp
  1             5             10             15
Val Leu Lys Ala Asp Gly Ala Ile Leu Val Asp Phe Trp Ala Glu Trp
             20             25             30
Cys Gly Pro Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp
             35             40             45
Glu Tyr Gln Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn
             50             55             60
Pro Gly Thr Ala Pro Lys Tyr Gly Ile Arg Gly Ile Pro Thr Leu Leu
             65             70             75
Leu Phe Lys Asn Gly Glu Val Ala Ala Thr Lys Val Gly Ala Leu Ser
             85             90             95
Lys Gly Gln Leu Lys Glu Phe Leu Asp Ala Asn Leu Ala
             100             105

```

<210> 44
 <211> 966
 <212> DNA
 <213> E. coli

<220>

<221> CDS

<222> (1)...(966)

<400> 44

atg ggc acg acc aaa cac agt aaa ctg ctt atc ctg ggt tca ggc ccg	48
Met Gly Thr Thr Lys His Ser Lys Leu Leu Ile Leu Gly Ser Gly Pro	
1 5 10 15	
gcg gga tac acc gct gct gtc tac gcg gcg cgc gcc aac ctg caa cct	96
Ala Gly Tyr Thr Ala Ala Val Tyr Ala Ala Arg Ala Asn Leu Gln Pro	
20 25 30	
gtg ctg att acc ggc atg gaa aaa ggc gcg caa ctg acc acc acc acg	144
Val Leu Ile Thr Gly Met Glu Lys Gly Gly Gln Leu Thr Thr Thr	
35 40 45	
gaa gtg gaa aac tgg cct ggc gat cca aac gat ctg acc ggt ccg tta	192
Glu Val Glu Asn Trp Pro Gly Asp Pro Asn Asp Leu Thr Gly Pro Leu	
50 55 60	
tta atg gag cgc atg cac gaa cat gcc acc aag ttt gaa act gag atc	240
Leu Met Glu Arg Met His Glu His Ala Thr Lys Phe Glu Thr Glu Ile	
65 70 75 80	
att ttt gat cat atc aac aag gtg gat ctg caa aac cgt ccg ttc cgt	288
Ile Phe Asp His Ile Asn Lys Val Asp Leu Gln Asn Arg Pro Phe Arg	
85 90 95	
ctg aat ggc gat aac ggc gaa tac act tgc gac gcg ctg att att gcc	336
Leu Asn Gly Asp Asn Gly Glu Tyr Thr Cys Asp Ala Leu Ile Ile Ala	
100 105 110	
acc gga gct tct gca cgc tat ctc ggc ctg ccc tct gaa gaa gcc ttt	384
Thr Gly Ala Ser Ala Arg Tyr Leu Gly Leu Pro Ser Glu Glu Ala Phe	
115 120 125	
aaa ggc cgt ggg gtt tct gct tgt gca acc tgc gac ggt ttc ttc tat	432
Lys Gly Arg Gly Val Ser Ala Cys Ala Thr Cys Asp Gly Phe Phe Tyr	
130 135 140	
cgc aac cag aaa gtt gcg gtc atc ggc gcg ggc aat acc gcg gtt gaa	480
Arg Asn Gln Lys Val Ala Val Ile Gly Gly Gly Asn Thr Ala Val Glu	
145 150 155 160	
gag gcg ttg tat ctg tct aac atc gct tgc gaa gtg cat ctg att cac	528
Glu Ala Leu Tyr Leu Ser Asn Ile Ala Ser Glu Val His Leu Ile His	
165 170 175	
cgc cgt gac ggt ttc cgc gcg gaa aaa atc ctc att aag cgc ctg atg	576
Arg Arg Asp Gly Phe Arg Ala Glu Lys Ile Leu Ile Lys Arg Leu Met	
180 185 190	
gat aaa gtg gag aac ggc aac atc att ctg cac acc aac cgt acg ctg	624
Asp Lys Val Glu Asn Gly Asn Ile Ile Leu His Thr Asn Arg Thr Leu	
195 200 205	
gaa gaa gtg acc ggc gat caa atg ggt gtc act ggc gtt cgt ctg cgc	672
Glu Glu Val Thr Gly Asp Gln Met Gly Val Thr Gly Val Arg Leu Arg	
210 215 220	
gat acg caa aac agc gat aac atc gag tca ctc gac gtt gcc ggt ctg	720
Asp Thr Gln Asn Ser Asp Asn Ile Glu Ser Leu Asp Val Ala Gly Leu	
225 230 235 240	
ttt gtt gct atc ggt cac agc ccg aat act gcg att ttc gaa ggg cag	768
Phe Val Ala Ile Gly His Ser Pro Asn Thr Ala Ile Phe Glu Gly Gln	

	245	250	255	
ctg gaa ctg gaa aac ggc tac atc aaa gta cag tcg ggt att cat ggt				816
Leu Glu Leu Glu Asn Gly Tyr Ile Lys Val Gln Ser Gly Ile His Gly				
	260	265	270	
aat gcc acc cag acc agc att cct ggc gtc ttt gcc gca ggc gac gtg				864
Asn Ala Thr Gln Thr Ser Ile Pro Gly Val Phe Ala Ala Gly Asp Val				
	275	280	285	
atg gat cac att tat cgc cag gcc att act tcg gcc ggt aca ggc tgc				912
Met Asp His Ile Tyr Arg Gln Ala Ile Thr Ser Ala Gly Thr Gly Cys				
	290	295	300	
atg gca gca ctt gat gcg gaa cgc tac ctc gat ggt tta gct gac gca				960
Met Ala Ala Leu Asp Ala Glu Arg Tyr Leu Asp Gly Leu Ala Asp Ala				
	305	310	315	320
aaa taa				966
Lys *				

<210> 45
 <211> 321
 <212> PRT
 <213> E. coli

<400> 45

Met Gly Thr Thr Lys His Ser Lys Leu Leu Ile Leu Gly Ser Gly Pro				
1 5 10 15				
Ala Gly Tyr Thr Ala Ala Val Tyr Ala Ala Arg Ala Asn Leu Gln Pro				
20 25 30				
Val Leu Ile Thr Gly Met Glu Lys Gly Gly Gln Leu Thr Thr Thr				
35 40 45				
Glu Val Glu Asn Trp Pro Gly Asp Pro Asn Asp Leu Thr Gly Pro Leu				
50 55 60				
Leu Met Glu Arg Met His Glu His Ala Thr Lys Phe Glu Thr Glu Ile				
65 70 75 80				
Ile Phe Asp His Ile Asn Lys Val Asp Leu Gln Asn Arg Pro Phe Arg				
85 90 95				
Leu Asn Gly Asp Asn Gly Glu Tyr Thr Cys Asp Ala Leu Ile Ile Ala				
100 105 110				
Thr Gly Ala Ser Ala Arg Tyr Leu Gly Leu Pro Ser Glu Glu Ala Phe				
115 120 125				
Lys Gly Arg Gly Val Ser Ala Cys Ala Thr Cys Asp Gly Phe Phe Tyr				
130 135 140				
Arg Asn Gln Lys Val Ala Val Ile Gly Gly Gly Asn Thr Ala Val Glu				
145 150 155 160				
Glu Ala Leu Tyr Leu Ser Asn Ile Ala Ser Glu Val His Leu Ile His				
165 170 175				
Arg Arg Asp Gly Phe Arg Ala Glu Lys Ile Leu Ile Lys Arg Leu Met				
180 185 190				
Asp Lys Val Glu Asn Gly Asn Ile Ile Leu His Thr Asn Arg Thr Leu				
195 200 205				
Glu Glu Val Thr Gly Asp Gln Met Gly Val Thr Gly Val Arg Leu Arg				
210 215 220				
Asp Thr Gln Asn Ser Asp Asn Ile Glu Ser Leu Asp Val Ala Gly Leu				
225 230 235 240				
Phe Val Ala Ile Gly His Ser Pro Asn Thr Ala Ile Phe Glu Gly Gln				
245 250 255				
Leu Glu Leu Glu Asn Gly Tyr Ile Lys Val Gln Ser Gly Ile His Gly				
260 265 270				
Asn Ala Thr Gln Thr Ser Ile Pro Gly Val Phe Ala Ala Gly Asp Val				
275 280 285				
Met Asp His Ile Tyr Arg Gln Ala Ile Thr Ser Ala Gly Thr Gly Cys				
290 295 300				

Met Ala Ala Leu Asp Ala Glu Arg Tyr Leu Asp Gly Leu Ala Asp Ala
 305 310 315 320
 Lys

<210> 46
 <211> 318
 <212> DNA
 <213> Homo Sapien

<220>
 <221> CDS
 <222> (1)...(318)

<400> 46
 atg gtg aag cag atc gag agc aag act gct ttt cag gaa gcc ttg gac 48
 Met Val Lys Gln Ile Glu Ser Lys Thr Ala Phe Gln Glu Ala Leu Asp
 1 5 10 15
 gct gca ggt gat aaa ctt gta gta gtt gac ttc tca gcc acg tgg tgt 96
 Ala Ala Gly Asp Lys Leu Val Val Val Asp Phe Ser Ala Thr Trp Cys
 20 25 30
 ggg cct tgc aaa atg atc aag cct ttc ttt cat tcc ctc tct gaa aag 144
 Gly Pro Cys Lys Met Ile Lys Pro Phe Phe His Ser Leu Ser Glu Lys
 35 40 45
 tat tcc aac gtg ata ttc ctt gaa gta gat gtg gat gac tgt cag gat 192
 Tyr Ser Asn Val Ile Phe Leu Glu Val Asp Val Asp Asp Cys Gln Asp
 50 55 60
 gtt gct tca gag tgt gaa gtc aaa tgc atg cca aca ttc cag ttt ttt 240
 Val Ala Ser Glu Cys Glu Val Lys Cys Met Pro Thr Phe Gln Phe Phe
 65 70 75 80
 aag aag gga caa aag gtg ggt gaa ttt tct gga gcc aat aag gaa aag 288
 Lys Lys Gly Gln Lys Val Gly Glu Phe Ser Gly Ala Asn Lys Glu Lys
 85 90 95
 ctt gaa gcc acc att aat gaa tta gtc taa 318
 Leu Glu Ala Thr Ile Asn Glu Leu Val *
 100 105

<210> 47
 <211> 105
 <212> PRT
 <213> Homo Sapien

<400> 47
 Met Val Lys Gln Ile Glu Ser Lys Thr Ala Phe Gln Glu Ala Leu Asp
 1 5 10 15
 Ala Ala Gly Asp Lys Leu Val Val Val Asp Phe Ser Ala Thr Trp Cys
 20 25 30
 Gly Pro Cys Lys Met Ile Lys Pro Phe Phe His Ser Leu Ser Glu Lys
 35 40 45
 Tyr Ser Asn Val Ile Phe Leu Glu Val Asp Val Asp Asp Cys Gln Asp
 50 55 60
 Val Ala Ser Glu Cys Glu Val Lys Cys Met Pro Thr Phe Gln Phe Phe
 65 70 75 80
 Lys Lys Gly Gln Lys Val Gly Glu Phe Ser Gly Ala Asn Lys Glu Lys
 85 90 95
 Leu Glu Ala Thr Ile Asn Glu Leu Val
 100 105

<210> 48
 <211> 1494
 <212> DNA
 <213> Homo sapien

<220>
 <221> CDS
 <222> (1)...(1494)

<400> 48

atg aac ggc cct gaa gat ctt ccc aag tcc tat gac tat gac ctt atc	48
Met Asn Gly Pro Glu Asp Leu Pro Lys Ser Tyr Asp Tyr Asp Leu Ile	
1 5 10 15	
atc att gga ggt ggc tca gga ggt ctg gca gct gct aag gag cca gcc	96
Ile Ile Gly Gly Gly Ser Gly Gly Leu Ala Ala Ala Lys Glu Pro Ala	
20 25 30	
caa tat ggc aag aag gtg atg gtc ctg gac ttt ggc act ccc acc cct	144
Gln Tyr Gly Lys Lys Val Met Val Leu Asp Phe Gly Thr Pro Thr Pro	
35 40 45	
ctt gga act aga tgg ggt ctt gga gga aca tgt gtg aat gtg ggt tgc	192
Leu Gly Thr Arg Trp Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys	
50 55 60	
ata cct aaa aaa ctg atg cat caa gca gct ttg tta gga caa gcc ctg	240
Ile Pro Lys Lys Leu Met His Gln Ala Ala Leu Leu Gly Gln Ala Leu	
65 70 75 80	
caa gac tct cga aat tat gga tgg aaa gtc gag gag aca gtt aag cat	288
Gln Asp Ser Arg Asn Tyr Gly Trp Lys Val Glu Glu Thr Val Lys His	
85 90 95	
gat tgg gac aga atg ata gaa gct gta cag aat cac att ggc tct ttg	336
Asp Trp Asp Arg Met Ile Glu Ala Val Gln Asn His Ile Gly Ser Leu	
100 105 110	
aat tgg ggc tac cga gta gct ctg cgg gag aaa aaa gtc gtc tat gag	384
Asn Trp Gly Tyr Arg Val Ala Leu Arg Glu Lys Lys Val Val Tyr Glu	
115 120 125	
aat gct tat ggg caa ttt att ggt cct cac agg att aag gca aca aat	432
Asn Ala Tyr Gly Gln Phe Ile Gly Pro His Arg Ile Lys Ala Thr Asn	
130 135 140	
aat aaa ggc aaa gaa aaa att tat tca gca gag aga ttt ctc att gcc	480
Asn Lys Gly Lys Glu Lys Ile Tyr Ser Ala Glu Arg Phe Leu Ile Ala	
145 150 155 160	
act ggt gaa aga cca cgt tac ttg ggc atc cct ggt gac aaa gaa tac	528
Thr Gly Glu Arg Pro Arg Tyr Leu Gly Ile Pro Gly Asp Lys Glu Tyr	
165 170 175	
tgc atc agc agt gat gat ctt ttc tcc ttg cct tac tgc ccg ggt aag	576
Cys Ile Ser Ser Asp Asp Leu Phe Ser Leu Pro Tyr Cys Pro Gly Lys	
180 185 190	
aca ctg gtt gtt gga gca tcc tat gtc gct ttg gag tgc gct gga ttt	624
Thr Leu Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe	
195 200 205	
ctt gct ggt att ggt tta gac gtc act gtt atg gtt agg tcc att ctt	672
Leu Ala Gly Ile Gly Leu Asp Val Thr Val Met Val Arg Ser Ile Leu	
210 215 220	
ctt aga gga ttt gac cag gac atg gcc aac aaa att ggt gaa cac atg	720

Leu 225	Arg	Gly	Phe	Asp	Gln 230	Asp	Met	Ala	Asn 235	Lys	Ile	Gly	Glu	His	Met 240	
gaa Glu	gaa Glu	cat His	ggc Gly	atc Ile	aag Lys	ttt Phe	ata Ile	aga Arg	cag Gln	ttc Phe	gta Val	cca Pro	att Ile	aaa Lys	gtt Val	768
gaa Glu	caa Gln	att Ile	gaa Glu	gca Ala	ggg Gly	aca Thr	cca Pro	ggc Gly	cga Arg	ctc Leu	aga Arg	gta Val	gta Val	gct Ala	cag Gln	816
tcc Ser	acc Thr	aat Asn	agt Ser	gag Glu	gaa Glu	atc Ile	att Ile	gaa Glu	gga Gly	gaa Glu	tat Tyr	aat Asn	acg Thr	gtg Val	atg Met	864
ctg Leu	gca Ala	ata Ile	gga Gly	aga Arg	gat Asp	gct Ala	tgc Cys	aca Thr	aga Arg	aaa Lys	att Ile	ggc Gly	tta Leu	gaa Glu	acc Thr	912
gta Val	ggg Gly	gtg Val	aag Lys	ata Ile	aat Asn	gaa Glu	aag Lys	act Thr	gga Gly	aaa Lys	ata Ile	cct Pro	gtc Val	aca Thr	gat Asp	960
gaa Glu	gaa Glu	cag Gln	acc Thr	aat Asn	gtg Val	cct Pro	tac Tyr	atc Ile	tat Tyr	gcc Ala	att Ile	ggc Gly	gat Asp	ata Ile	ttg Leu	1008
gag Glu	gat Asp	aag Lys	gtg Val	gag Glu	ctc Leu	acc Thr	cca Pro	gtt Val	gca Ala	atc Ile	cag Gln	gca Ala	gga Gly	aga Arg	ttg Leu	1056
ctg Leu	gct Ala	cag Gln	agg Arg	ctc Leu	tat Tyr	gca Ala	ggt Gly	tcc Ser	act Thr	gtc Val	aag Lys	tgt Cys	gac Asp	tat Tyr	gaa Glu	1104
aat Asn	gtt Val	cca Pro	acc Thr	act Thr	gta Val	ttt Phe	act Thr	cct Pro	ttg Leu	gaa Glu	tat Tyr	ggt Gly	gct Ala	tgt Cys	ggc Gly	1152
ctt Leu	tct Ser	gag Glu	gag Glu	aaa Lys	gct Ala	gtg Val	gag Glu	aag Lys	ttt Phe	ggg Gly	gaa Glu	gaa Glu	aat Asn	att Ile	gag Glu	1200
gtt Val	tac Tyr	cat His	agt Ser	tac Tyr	ttt Phe	tgg Trp	cca Pro	ttg Leu	gaa Glu	tgg Trp	acg Thr	att Ile	ccg Pro	tca Ser	aga Arg	1248
gat Asp	aac Asn	aac Asn	aaa Lys	tgt Cys	tat Tyr	gca Ala	aaa Lys	ata Ile	atc Ile	tgt Cys	aat Asn	act Thr	aaa Lys	gac Asp	aat Asn	1296
gaa Glu	cgt Arg	gtt Val	gtg Val	ggc Gly	ttt Phe	cac His	gta Val	ctg Leu	ggt Gly	cca Pro	aat Asn	gct Ala	gga Gly	gaa Glu	gtt Val	1344
aca Thr	caa Gln	ggc Gly	ttt Phe	gca Ala	gct Ala	gcg Ala	ctc Leu	aaa Lys	tgt Cys	gga Gly	ctg Leu	acc Thr	aaa Lys	aag Lys	cag Gln	1392
ctg Leu	gac Asp	agc Ser	aca Thr	att Ile	gga Gly	atc Ile	cac His	cct Pro	gtc Val	tgt Cys	gca Ala	gag Glu	gta Val	ttc Phe	aca Thr	1440
aca Thr	ttg Leu	tct Ser	gtg Val	acc Thr	aag Lys	cgc Arg	tct Ser	ggg Gly	gca Ala	agc Ser	atc Ile	ctc Leu	cag Gln	gct Ala	ggc Gly	1488

tgc tga
Cys *

1494

<210> 49
<211> 497
<212> PRT
<213> Homo sapien

<400> 49

Met	Asn	Gly	Pro	Glu	Asp	Leu	Pro	Lys	Ser	Tyr	Asp	Tyr	Asp	Leu	Ile
1				5				10						15	
Ile	Ile	Gly	Gly	Gly	Ser	Gly	Gly	Leu	Ala	Ala	Ala	Lys	Glu	Pro	Ala
		20						25					30		
Gln	Tyr	Gly	Lys	Lys	Val	Met	Val	Leu	Asp	Phe	Gly	Thr	Pro	Thr	Pro
		35					40					45			
Leu	Gly	Thr	Arg	Trp	Gly	Leu	Gly	Gly	Thr	Cys	Val	Asn	Val	Gly	Cys
50						55					60				
Ile	Pro	Lys	Lys	Leu	Met	His	Gln	Ala	Ala	Leu	Leu	Gly	Gln	Ala	Leu
65					70					75					80
Gln	Asp	Ser	Arg	Asn	Tyr	Gly	Trp	Lys	Val	Glu	Glu	Thr	Val	Lys	His
				85					90					95	
Asp	Trp	Asp	Arg	Met	Ile	Glu	Ala	Val	Gln	Asn	His	Ile	Gly	Ser	Leu
				100					105					110	
Asn	Trp	Gly	Tyr	Arg	Val	Ala	Leu	Arg	Glu	Lys	Lys	Val	Val	Tyr	Glu
		115					120						125		
Asn	Ala	Tyr	Gly	Gln	Phe	Ile	Gly	Pro	His	Arg	Ile	Lys	Ala	Thr	Asn
		130				135					140				
Asn	Lys	Gly	Lys	Glu	Lys	Ile	Tyr	Ser	Ala	Glu	Arg	Phe	Leu	Ile	Ala
145					150					155					160
Thr	Gly	Glu	Arg	Pro	Arg	Tyr	Leu	Gly	Ile	Pro	Gly	Asp	Lys	Glu	Tyr
				165					170					175	
Cys	Ile	Ser	Ser	Asp	Asp	Leu	Phe	Ser	Leu	Pro	Tyr	Cys	Pro	Gly	Lys
			180					185						190	
Thr	Leu	Val	Val	Gly	Ala	Ser	Tyr	Val	Ala	Leu	Glu	Cys	Ala	Gly	Phe
		195					200					205			
Leu	Ala	Gly	Ile	Gly	Leu	Asp	Val	Thr	Val	Met	Val	Arg	Ser	Ile	Leu
		210				215					220				
Leu	Arg	Gly	Phe	Asp	Gln	Asp	Met	Ala	Asn	Lys	Ile	Gly	Glu	His	Met
225					230					235					240
Glu	Glu	His	Gly	Ile	Lys	Phe	Ile	Arg	Gln	Phe	Val	Pro	Ile	Lys	Val
				245					250					255	
Glu	Gln	Ile	Glu	Ala	Gly	Thr	Pro	Gly	Arg	Leu	Arg	Val	Val	Ala	Gln
			260					265					270		
Ser	Thr	Asn	Ser	Glu	Glu	Ile	Ile	Glu	Gly	Glu	Tyr	Asn	Thr	Val	Met
		275					280					285			
Leu	Ala	Ile	Gly	Arg	Asp	Ala	Cys	Thr	Arg	Lys	Ile	Gly	Leu	Glu	Thr
		290				295					300				
Val	Gly	Val	Lys	Ile	Asn	Glu	Lys	Thr	Gly	Lys	Ile	Pro	Val	Thr	Asp
305					310					315					320
Glu	Glu	Gln	Thr	Asn	Val	Pro	Tyr	Ile	Tyr	Ala	Ile	Gly	Asp	Ile	Leu
				325					330					335	
Glu	Asp	Lys	Val	Glu	Leu	Thr	Pro	Val	Ala	Ile	Gln	Ala	Gly	Arg	Leu
			340					345					350		
Leu	Ala	Gln	Arg	Leu	Tyr	Ala	Gly	Ser	Thr	Val	Lys	Cys	Asp	Tyr	Glu
		355					360					365			
Asn	Val	Pro	Thr	Thr	Val	Phe	Thr	Pro	Leu	Glu	Tyr	Gly	Ala	Cys	Gly
		370				375					380				
Leu	Ser	Glu	Glu	Lys	Ala	Val	Glu	Lys	Phe	Gly	Glu	Glu	Asn	Ile	Glu
385					390					395					400
Val	Tyr	His	Ser	Tyr	Phe	Trp	Pro	Leu	Glu	Trp	Thr	Ile	Pro	Ser	Arg
				405					410					415	
Asp	Asn	Asn	Lys	Cys	Tyr	Ala	Lys	Ile	Cys	Asn	Thr	Lys	Asp	Asn	
			420					425				430			
Glu	Arg	Val	Val	Gly	Phe	His	Val	Leu	Gly	Pro	Asn	Ala	Gly	Glu	Val
		435					440					445			

Thr Gln Gly Phe Ala Ala Ala Leu Lys Cys Gly Leu Thr Lys Lys Gln
 450 455 460
 Leu Asp Ser Thr Ile Gly Ile His Pro Val Cys Ala Glu Val Phe Thr
 465 470 475 480
 Thr Leu Ser Val Thr Lys Arg Ser Gly Ala Ser Ile Leu Gln Ala Gly
 485 490 495
 Cys

<210> 50
 <211> 1377
 <212> DNA
 <213> Mycobacterium leprae

<220>
 <221> CDS
 <222> (1)...(1377)

<400> 50
 atg aac acc act cct tct gcg cat gag acg ata cac gaa gtg atc gtt 48
 Met Asn Thr Thr Pro Ser Ala His Glu Thr Ile His Glu Val Ile Val
 1 5 10 15
 att ggc tcc ggt cca gca ggc tac act gct gcc ctg tac gcc gct cgt 96
 Ile Gly Ser Gly Pro Ala Gly Tyr Thr Ala Ala Leu Tyr Ala Ala Arg
 20 25 30
 gca cag cta aca ccg ctg gta ttt gag ggt acc tca ttc ggc ggc gcg 144
 Ala Gln Leu Thr Pro Leu Val Phe Glu Gly Thr Ser Phe Gly Gly Ala
 35 40 45
 ctg atg acc acc acc gag gtg gaa aac tac cca ggt ttt cgc aac ggc 192
 Leu Met Thr Thr Thr Glu Val Glu Asn Tyr Pro Gly Phe Arg Asn Gly
 50 55 60
 ata acc ggc ccg gag ttg atg gac gat atg cgt gaa cag gca ctg cga 240
 Ile Thr Gly Pro Glu Leu Met Asp Asp Met Arg Glu Gln Ala Leu Arg
 65 70 75 80
 ttc ggc gcg gaa ctg cgg acc gaa gac gtc gag tcg gta tca ttg cgt 288
 Phe Gly Ala Glu Leu Arg Thr Glu Asp Val Glu Ser Val Ser Leu Arg
 85 90 95
 ggc ccg atc aaa tcg gtc gtc acc gct gaa gga cag act tat cag gcc 336
 Gly Pro Ile Lys Ser Val Val Thr Ala Glu Gly Gln Thr Tyr Gln Ala
 100 105 110
 cga gcc gtc atc ctc gcc atg ggt acc tcc gtg cgt tat cta cag atc 384
 Arg Ala Val Ile Leu Ala Met Gly Thr Ser Val Arg Tyr Leu Gln Ile
 115 120 125
 ccc ggc gag caa gaa ttg cta gga cgt ggc gtg agt gca tgc gcg acc 432
 Pro Gly Glu Gln Glu Leu Leu Gly Arg Gly Val Ser Ala Cys Ala Thr
 130 135 140
 tgc gac ggg tcc ttt ttc cgc ggc caa gac att gcc gtc att ggc ggt 480
 Cys Asp Gly Ser Phe Phe Arg Gly Gln Asp Ile Ala Val Ile Gly Gly
 145 150 155 160
 gga gac tca gcg atg gag gaa gcc ctc ttt ttg acc cgg ttc gcc cgc 528
 Gly Asp Ser Ala Met Glu Glu Ala Leu Phe Leu Thr Arg Phe Ala Arg
 165 170 175
 agc gtc acg ctc gtg cac cgc cgc gac gaa ttc cga gct tct aag atc 576
 Ser Val Thr Leu Val His Arg Arg Asp Glu Phe Arg Ala Ser Lys Ile
 180 185 190

atg ctc ggt cgc gcc cgt aac aat gac aag atc aaa ttc atc acc aac	624
Met Leu Gly Arg Ala Arg Asn Asn Asp Lys Ile Lys Phe Ile Thr Asn	
195 200 205	
cac acc gtg gtc gcg gtg aac ggg tat aca aca gtg acc gga ttg cgg	672
His Thr Val Val Ala Val Asn Gly Tyr Thr Thr Val Thr Gly Leu Arg	
210 215 220	
ttg cgt aac acc aca acg gga gag gaa acc acg cta gta gtg acc ggg	720
Leu Arg Asn Thr Thr Thr Gly Glu Glu Thr Thr Leu Val Val Thr Gly	
225 230 235 240	
gtt ttt gtt gca att ggc cat gaa cca cgt tcc agc ctg gtg agc gat	768
Val Phe Val Ala Ile Gly His Glu Pro Arg Ser Ser Leu Val Ser Asp	
245 250 255	
gtc gtc gac ata gac ccg gat ggc tac gtc ctg gtg aaa gga cgt acg	816
Val Val Asp Ile Asp Pro Asp Gly Tyr Val Leu Val Lys Gly Arg Thr	
260 265 270	
acg agt aca tcg atg gac ggc gtt ttt gcg gcc ggc gac ctg gta gat	864
Thr Ser Thr Ser Met Asp Gly Val Phe Ala Ala Gly Asp Leu Val Asp	
275 280 285	
cgc acc tac cgg cag gcg atc act gcc gca ggt agt ggc tgt gcc gcc	912
Arg Thr Tyr Arg Gln Ala Ile Thr Ala Ala Gly Ser Gly Cys Ala Ala	
290 295 300	
gcc atc gac gcc gaa cgt tgg ttg gcg gag cat gcc ggg tca aaa gct	960
Ala Ile Asp Ala Glu Arg Trp Leu Ala Glu His Ala Gly Ser Lys Ala	
305 310 315 320	
aac gaa aca aca gag gaa act gga gac gtt gac agt acc gac aca acc	1008
Asn Glu Thr Thr Glu Glu Thr Gly Asp Val Asp Ser Thr Asp Thr Thr	
325 330 335	
gat tgg agc act gcg atg act gac gcc aag aac gcc ggg gtc aca ata	1056
Asp Trp Ser Thr Ala Met Thr Asp Ala Lys Asn Ala Gly Val Thr Ile	
340 345 350	
gaa gtg acc gat gct tcc ttt ttc gca gac gtc tta tcc agt aat aag	1104
Glu Val Thr Asp Ala Ser Phe Phe Ala Asp Val Leu Ser Ser Asn Lys	
355 360 365	
cct gtg tta gtt gat ttt tgg gca aca tgg tgt gga ccc tgc aag atg	1152
Pro Val Leu Val Asp Phe Trp Ala Thr Trp Cys Gly Pro Cys Lys Met	
370 375 380	
gta gcg ccg gta ctc gaa gag atc gcg tcc gaa caa cga aac cag ctc	1200
Val Ala Pro Val Leu Glu Glu Ile Ala Ser Glu Gln Arg Asn Gln Leu	
385 390 395 400	
act gtc gcc aag tta gat gta gac acc aac ccg gaa atg gca cgc gag	1248
Thr Val Ala Lys Leu Asp Val Asp Thr Asn Pro Glu Met Ala Arg Glu	
405 410 415	
ttc cag gtc gtg tcg ata ccc aca atg att ctg ttc cag ggt ggc caa	1296
Phe Gln Val Val Ser Ile Pro Thr Met Ile Leu Phe Gln Gly Gly Gln	
420 425 430	
cca gta aaa cgc atc gtt ggc gct aag ggc aaa gca gcg tta cta cgt	1344
Pro Val Lys Arg Ile Val Gly Ala Lys Gly Lys Ala Ala Leu Leu Arg	
435 440 445	
gac ctt tcc gac gtg gta cct aac ctc aat tag	1377
Asp Leu Ser Asp Val Val Pro Asn Leu Asn *	

450

455

<210> 51
 <211> 458
 <212> PRT
 <213> Mycobacterium leprae

<400> 51

Met	Asn	Thr	Thr	Pro	Ser	Ala	His	Glu	Thr	Ile	His	Glu	Val	Ile	Val
1				5					10				15		
Ile	Gly	Ser	Gly	Pro	Ala	Gly	Tyr	Thr	Ala	Ala	Leu	Tyr	Ala	Ala	Arg
			20					25					30		
Ala	Gln	Leu	Thr	Pro	Leu	Val	Phe	Glu	Gly	Thr	Ser	Phe	Gly	Gly	Ala
		35					40					45			
Leu	Met	Thr	Thr	Thr	Glu	Val	Glu	Asn	Tyr	Pro	Gly	Phe	Arg	Asn	Gly
	50				55						60				
Ile	Thr	Gly	Pro	Glu	Leu	Met	Asp	Asp	Met	Arg	Glu	Gln	Ala	Leu	Arg
65					70					75				80	
Phe	Gly	Ala	Glu	Leu	Arg	Thr	Glu	Asp	Val	Glu	Ser	Val	Ser	Leu	Arg
				85				90						95	
Gly	Pro	Ile	Lys	Ser	Val	Val	Thr	Ala	Glu	Gly	Gln	Thr	Tyr	Gln	Ala
			100					105						110	
Arg	Ala	Val	Ile	Leu	Ala	Met	Gly	Thr	Ser	Val	Arg	Tyr	Leu	Gln	Ile
		115					120					125			
Pro	Gly	Glu	Gln	Glu	Leu	Leu	Gly	Arg	Gly	Val	Ser	Ala	Cys	Ala	Thr
	130					135					140				
Cys	Asp	Gly	Ser	Phe	Phe	Arg	Gly	Gln	Asp	Ile	Ala	Val	Ile	Gly	Gly
145					150					155				160	
Gly	Asp	Ser	Ala	Met	Glu	Glu	Ala	Leu	Phe	Leu	Thr	Arg	Phe	Ala	Arg
				165				170						175	
Ser	Val	Thr	Leu	Val	His	Arg	Arg	Asp	Glu	Phe	Arg	Ala	Ser	Lys	Ile
			180					185					190		
Met	Leu	Gly	Arg	Ala	Arg	Asn	Asn	Asp	Lys	Ile	Lys	Phe	Ile	Thr	Asn
		195					200					205			
His	Thr	Val	Val	Ala	Val	Asn	Gly	Tyr	Thr	Thr	Val	Thr	Gly	Leu	Arg
	210					215					220				
Leu	Arg	Asn	Thr	Thr	Thr	Gly	Glu	Glu	Thr	Thr	Leu	Val	Val	Thr	Gly
225					230					235				240	
Val	Phe	Val	Ala	Ile	Gly	His	Glu	Pro	Arg	Ser	Ser	Leu	Val	Ser	Asp
				245				250						255	
Val	Val	Asp	Ile	Asp	Pro	Asp	Gly	Tyr	Val	Leu	Val	Lys	Gly	Arg	Thr
		260					265						270		
Thr	Ser	Thr	Ser	Met	Asp	Gly	Val	Phe	Ala	Ala	Gly	Asp	Leu	Val	Asp
		275					280					285			
Arg	Thr	Tyr	Arg	Gln	Ala	Ile	Thr	Ala	Ala	Gly	Ser	Gly	Cys	Ala	Ala
	290					295					300				
Ala	Ile	Asp	Ala	Glu	Arg	Trp	Leu	Ala	Glu	His	Ala	Gly	Ser	Lys	Ala
305					310					315				320	
Asn	Glu	Thr	Thr	Glu	Glu	Thr	Gly	Asp	Val	Asp	Ser	Thr	Asp	Thr	Thr
				325				330						335	
Asp	Trp	Ser	Thr	Ala	Met	Thr	Asp	Ala	Lys	Asn	Ala	Gly	Val	Thr	Ile
		340						345					350		
Glu	Val	Thr	Asp	Ala	Ser	Phe	Phe	Ala	Asp	Val	Leu	Ser	Ser	Asn	Lys
		355					360					365			
Pro	Val	Leu	Val	Asp	Phe	Trp	Ala	Thr	Trp	Cys	Gly	Pro	Cys	Lys	Met
	370					375					380				
Val	Ala	Pro	Val	Leu	Glu	Glu	Ile	Ala	Ser	Glu	Gln	Arg	Asn	Gln	Leu
385					390					395				400	
Thr	Val	Ala	Lys	Leu	Asp	Val	Asp	Thr	Asn	Pro	Glu	Met	Ala	Arg	Glu
				405				410						415	
Phe	Gln	Val	Val	Ser	Ile	Pro	Thr	Met	Ile	Leu	Phe	Gln	Gly	Gly	Gln
		420						425					430		
Pro	Val	Lys	Arg	Ile	Val	Gly	Ala	Lys	Gly	Lys	Ala	Ala	Leu	Leu	Arg
		435					440					445			
Asp	Leu	Ser	Asp	Val	Val	Pro	Asn	Leu	Asn						
	450					455									

<210> 52
 <211> 178
 <212> PRT
 <213> Arabidopsis thaliana

<400> 52
 Met Pro Leu Ser Leu Arg Leu Ser Pro Ser Pro Thr Ala Leu Ser Pro
 1 5 10 15
 Thr Thr Gly Gly Phe Gly Pro Ser Arg Lys Gln Cys Arg Ile Pro Tyr
 20 25 30
 Ser Gly Val Pro Thr Thr Lys Ile Gly Phe Cys Ser Leu Asp Ser Arg
 35 40 45
 Lys Arg Gly Asp Ser Ser Val Val Arg Cys Ser Leu Glu Thr Val Asn
 50 55 60
 Val Ser Val Gly Gln Val Thr Glu Val Asp Lys Asp Thr Phe Trp Pro
 65 70 75 80
 Ile Val Lys Ala Ala Gly Glu Lys Leu Val Val Leu Asp Met Tyr Thr
 85 90 95
 Gln Trp Cys Gly Pro Cys Lys Val Ile Ala Pro Lys Tyr Lys Ala Leu
 100 105 110
 Ser Glu Lys Tyr Asp Asp Val Val Phe Leu Lys Leu Asp Cys Asn Pro
 115 120 125
 Asp Asn Arg Pro Leu Pro Lys Glu Leu Gly Ile Arg Val Val Pro Thr
 130 135 140
 Phe Lys Ile Leu Lys Asp Asn Lys Val Val Lys Glu Val Thr Gly Ala
 145 150 155 160
 Lys Tyr Asp Asp Leu Val Ala Ala Ile Glu Thr Ala Arg Ser Ala Ala
 165 170 175
 Ser Gly

<210> 53
 <211> 185
 <212> PRT
 <213> Arabidopsis thaliana

<400> 53
 Met Pro Leu Ser Leu Arg Leu Ala Pro Ser Pro Thr Ser Phe Arg Tyr
 1 5 10 15
 Ser Pro Ile Thr Ser Thr Gly Ala Gly Gly Phe Ser Pro Val Lys Gln
 20 25 30
 His Cys Arg Ile Pro Asn Ser Gly Val Ala Thr Lys Ile Gly Phe Cys
 35 40 45
 Ser Gly Gly Gly Val Leu Asp Ser Gly Arg Arg Ile Gly Ser Cys
 50 55 60
 Val Val Arg Cys Ser Leu Glu Thr Val Asn Val Thr Val Gly Gln Val
 65 70 75 80
 Thr Glu Val Asp Lys Asp Thr Phe Trp Pro Ile Val Lys Ala Ala Gly
 85 90 95
 Asp Lys Ile Val Val Leu Asp Met Tyr Thr Gln Trp Cys Gly Pro Cys
 100 105 110
 Lys Val Ile Ala Pro Lys Tyr Lys Glu Leu Ser Glu Lys Tyr Gln Asp
 115 120 125
 Met Val Phe Leu Lys Leu Asp Cys Asn Gln Asp Asn Lys Pro Leu Ala
 130 135 140
 Lys Glu Leu Gly Ile Arg Val Val Pro Thr Phe Lys Ile Leu Lys Asp
 145 150 155 160
 Asn Lys Val Val Lys Glu Val Thr Gly Ala Lys Tyr Glu Asp Leu Leu
 165 170 175
 Ala Ala Ile Glu Ala Ala Arg Ser Gly
 180 185

<210> 54
 <211> 182

<212> PRT

<213> Brassica napus

<400> 54

```

Met Pro Leu Ser Leu Arg Leu Ala Pro Ser Pro Thr Ala Leu Ser Pro
 1          5          10          15
Thr Thr Gly Gly Phe Ser Pro Ala Lys Lys Gln Cys Arg Ile Pro Ser
      20          25          30
Tyr Ser Gly Val Ala Thr Thr Thr Arg Arg Ile Gly Leu Cys Ser Leu
      35          40          45
Asp Tyr Val Lys Arg Gly Asp Ser Ser Val Val Arg Cys Ser Leu Gln
      50          55          60
Thr Val Asn Val Ser Val Gly Gln Val Thr Glu Val Asp Lys Asp Thr
      65          70          75          80
Phe Trp Pro Ile Val Lys Ala Ala Gly Glu Lys Ile Val Val Leu Asp
      85          90          95
Met Tyr Thr Gln Trp Cys Gly Pro Cys Lys Val Ile Ala Pro Lys Tyr
      100          105          110
Lys Ala Leu Ser Glu Lys Tyr Glu Asp Val Val Phe Leu Lys Leu Asp
      115          120          125
Cys Asn Pro Glu Asn Arg Pro Leu Ala Lys Glu Leu Gly Ile Arg Val
      130          135          140
Val Pro Thr Phe Lys Ile Leu Lys Asp Asn Gln Val Val Lys Glu Val
      145          150          155          160
Thr Gly Ala Lys Tyr Asp Asp Leu Val Ala Ala Ile Glu Thr Ala Arg
      165          170          175
Ser Ala Ser Ser Ser Gly
      180

```

<210> 55

<211> 191

<212> PRT

<213> Mesembryanthemum crystallinum

<400> 55

```

Met Ala Met Gln Leu Ser Leu Ser His Gln Ser Trp Ala Lys Ser Leu
 1          5          10          15
Ala Ser Pro Ile Thr Ser Phe Asp Pro Ala Arg Ser Pro Pro Lys Arg
      20          25          30
Val Glu Leu Gly Pro Asn Cys Leu Asn Gly Gly Ala Thr Ala Gly Lys
      35          40          45
Leu Met Arg Glu Lys Val Gly Glu Arg Met Arg Met Ser Gly Arg Ser
      50          55          60
Cys Cys Val Lys Ala Ser Leu Glu Thr Ala Val Gly Ala Glu Ser Glu
      65          70          75          80
Thr Leu Val Gly Lys Val Thr Glu Val Asp Lys Asp Thr Phe Trp Pro
      85          90          95
Ile Ala Asn Gly Ala Gly Asp Lys Pro Val Val Leu Asp Met Tyr Thr
      100          105          110
Gln Trp Cys Gly Pro Cys Lys Val Met Ala Pro Lys Tyr Gln Glu Leu
      115          120          125
Ala Glu Lys Leu Leu Asp Val Val Phe Leu Lys Leu Asp Cys Asn Gln
      130          135          140
Glu Asn Lys Pro Leu Ala Lys Glu Leu Gly Ile Arg Val Val Pro Thr
      145          150          155          160
Phe Lys Ile Leu Lys Gly Gly Lys Ile Val Asp Glu Val Thr Gly Ala
      165          170          175
Lys Phe Asp Lys Leu Val Ala Ala Ile Glu Ala Ala Arg Ser Ser
      180          185          190

```

<210> 56

<211> 182

<212> PRT

<213> Pisum sativum

<400> 56

```

Met Ala Leu Asn Leu Cys Thr Ser Pro Lys Trp Ile Gly Thr Thr Val
 1          5          10          15
Phe Asp Ser Ala Ser Ser Ser Lys Pro Ser Leu Ala Ser Ser Phe Ser
          20          25          30
Thr Thr Ser Phe Ser Ser Ser Ile Leu Cys Ser Lys Arg Val Gly Leu
          35          40          45
Gln Arg Leu Ser Leu Arg Arg Ser Ile Ser Val Ser Val Arg Ser Ser
          50          55          60
Leu Glu Thr Ala Gly Pro Thr Val Thr Val Gly Lys Val Thr Glu Val
65          70          75          80
Asn Lys Asp Thr Phe Trp Pro Ile Val Asn Ala Ala Gly Asp Lys Thr
          85          90          95
Val Val Leu Asp Met Phe Thr Lys Trp Cys Gly Pro Cys Lys Val Ile
          100          105          110
Ala Pro Leu Tyr Glu Glu Leu Ser Gln Lys Tyr Leu Asp Val Val Phe
          115          120          125
Leu Lys Leu Asp Cys Asn Gln Asp Asn Lys Ser Leu Ala Lys Glu Leu
          130          135          140
Gly Ile Lys Val Val Pro Thr Phe Lys Ile Leu Lys Asp Asn Lys Ile
145          150          155          160
Val Lys Glu Val Thr Gly Ala Lys Phe Asp Asp Leu Val Ala Ala Ile
          165          170          175
Asp Thr Val Arg Ser Ser
          180

```

<210> 57

<211> 190

<212> PRT

<213> Spinacia oleracea

<400> 57

```

Met Ala Leu His Leu Ser Leu Ser His Gln Ser Trp Thr Ser Pro Ala
 1          5          10          15
His Pro Ile Thr Ser Ser Asp Pro Thr Arg Ser Ser Val Pro Gly Thr
          20          25          30
Gly Leu Ser Arg Arg Val Asp Phe Leu Gly Ser Cys Lys Ile Asn Gly
          35          40          45
Val Phe Val Val Lys Arg Lys Asp Arg Arg Arg Met Arg Gly Gly Glu
50          55          60
Val Arg Ala Ser Met Glu Gln Ala Leu Gly Thr Gln Glu Met Glu Ala
65          70          75          80
Ile Val Gly Lys Val Thr Glu Val Asn Lys Asp Thr Phe Trp Pro Ile
          85          90          95
Val Lys Ala Ala Gly Asp Lys Pro Val Val Leu Asp Met Phe Thr Gln
          100          105          110
Trp Cys Gly Pro Cys Lys Ala Met Ala Pro Lys Tyr Glu Lys Leu Ala
          115          120          125
Glu Glu Tyr Leu Asp Val Ile Phe Leu Lys Leu Asp Cys Asn Gln Glu
130          135          140
Asn Lys Thr Leu Ala Lys Glu Leu Gly Ile Arg Val Val Pro Thr Phe
145          150          155          160
Lys Ile Leu Lys Glu Asn Ser Val Val Gly Glu Val Thr Gly Ala Lys
          165          170          175
Tyr Asp Lys Leu Leu Glu Ala Ile Gln Ala Ala Arg Ser Ser
          180          185          190

```

<210> 58

<211> 106

<212> PRT

<213> Anabaena

<400> 58

```

Ser Ala Ala Ala Gln Val Thr Asp Ser Thr Phe Lys Gln Glu Val Leu
 1          5          10          15

```

Asp Ser Asp Val Pro Val Leu Val Asp Phe Trp Ala Pro Trp Cys Gly
 20 25 30
 Pro Cys Arg Met Val Ala Pro Val Val Asp Glu Ile Ala Gln Gln Tyr
 35 40 45
 Glu Gly Lys Ile Lys Val Val Lys Val Asn Thr Asp Glu Asn Pro Gln
 50 55 60
 Val Ala Ser Gln Tyr Gly Ile Arg Ser Ile Pro Thr Leu Met Ile Phe
 65 70 75 80
 Lys Gly Gly Gln Lys Val Asp Met Val Val Gly Ala Val Pro Lys Thr
 85 90 95
 Thr Leu Ser Gln Thr Leu Glu Lys His Leu
 100 105

<210> 59

<211> 179

<212> PRT

<213> Arabidopsis thaliana

<400> 59

Met Ala Ala Tyr Thr Cys Thr Ser Arg Pro Pro Ile Ser Ile Arg Ser
 1 5 10 15
 Glu Met Arg Ile Ala Ser Ser Pro Thr Gly Ser Phe Ser Thr Arg Gln
 20 25 30
 Met Phe Ser Val Leu Pro Glu Ser Ser Gly Leu Arg Thr Arg Val Ser
 35 40 45
 Leu Ser Ser Leu Ser Lys Asn Ser Arg Val Ser Arg Leu Arg Arg Gly
 50 55 60
 Val Ile Cys Glu Ala Gln Asp Thr Ala Thr Gly Ile Pro Val Val Asn
 65 70 75 80
 Asp Ser Thr Trp Asp Ser Leu Val Leu Lys Ala Asp Glu Pro Val Phe
 85 90 95
 Val Asp Phe Trp Ala Pro Trp Cys Gly Pro Cys Lys Met Ile Asp Pro
 100 105 110
 Ile Val Asn Glu Leu Ala Gln Lys Tyr Ala Gly Gln Phe Lys Phe Tyr
 115 120 125
 Lys Leu Asn Thr Asp Glu Ser Pro Ala Thr Pro Gly Gln Tyr Gly Val
 130 135 140
 Arg Ser Ile Pro Thr Ile Met Ile Phe Val Asn Gly Glu Lys Lys Asp
 145 150 155 160
 Thr Ile Ile Gly Ala Val Ser Lys Asp Thr Leu Ala Thr Ser Ile Asn
 165 170 175
 Lys Phe Leu

<210> 60

<211> 186

<212> PRT

<213> Arabidopsis thaliana

<400> 60

Met Ala Ala Phe Thr Cys Thr Ser Arg Pro Pro Ile Ser Leu Arg Ser
 1 5 10 15
 Glu Thr Arg Ile Val Ser Ser Ser Pro Ser Ala Ser Ser Leu Ser Ser
 20 25 30
 Arg Arg Met Phe Ala Val Leu Pro Glu Ser Ser Gly Leu Arg Ile Arg
 35 40 45
 Leu Ser Leu Ser Pro Ala Ser Leu Thr Ser Ile His Gln Pro Arg Val
 50 55 60
 Ser Arg Leu Arg Arg Ala Val Val Cys Glu Ala Gln Glu Thr Thr Thr
 65 70 75 80
 Asp Ile Gln Val Val Asn Asp Ser Thr Trp Asp Ser Leu Val Leu Lys
 85 90 95
 Ala Thr Gly Pro Val Val Val Asp Phe Trp Ala Pro Trp Cys Gly Pro
 100 105 110
 Cys Lys Met Ile Asp Pro Leu Val Asn Asp Leu Ala Gln His Tyr Thr

		115					120					125					
Gly	Lys	Ile	Lys	Phe	Tyr	Lys	Leu	Asn	Thr	Asp	Glu	Ser	Pro	Asn	Thr		
	130					135					140						
Pro	Gly	Gln	Tyr	Gly	Val	Arg	Ser	Ile	Pro	Thr	Ile	Met	Ile	Phe	Val		
145					150					155					160		
Gly	Gly	Glu	Lys	Lys	Asp	Thr	Ile	Ile	Gly	Ala	Val	Pro	Lys	Thr	Thr		
			165						170					175			
Leu	Thr	Ser	Ser	Leu	Asp	Lys	Phe	Leu	Pro								
			180					185									

<210> 61

<211> 173

<212> PRT

<213> Arabidopsis thaliana

<400> 61

Met	Ala	Ile	Ser	Ser	Ser	Ser	Ser	Ser	Ile	Cys	Phe	Asn	Pro	Thr	Arg		
1				5					10					15			
Phe	His	Thr	Ala	Arg	His	Ile	Ser	Ser	Pro	Ser	Arg	Leu	Phe	Pro	Val		
			20					25					30				
Thr	Ser	Phe	Ser	Pro	Arg	Ser	Leu	Arg	Phe	Ser	Asp	Arg	Arg	Ser	Leu		
		35					40					45					
Leu	Ser	Ser	Ser	Ala	Ser	Arg	Leu	Arg	Leu	Ser	Pro	Leu	Cys	Val	Arg		
	50					55					60						
Asp	Ser	Arg	Ala	Ala	Glu	Val	Thr	Gln	Arg	Ser	Trp	Glu	Asp	Ser	Val		
65					70					75					80		
Leu	Lys	Ser	Glu	Thr	Pro	Val	Leu	Val	Glu	Phe	Tyr	Thr	Ser	Trp	Cys		
				85					90					95			
Gly	Pro	Cys	Arg	Met	Val	His	Arg	Ile	Ile	Asp	Glu	Ile	Ala	Gly	Asp		
			100					105					110				
Tyr	Ala	Gly	Lys	Leu	Asn	Cys	Tyr	Leu	Leu	Asn	Ala	Asp	Asn	Asp	Leu		
	115						120					125					
Pro	Val	Ala	Glu	Glu	Tyr	Glu	Ile	Lys	Ala	Val	Pro	Val	Val	Leu	Leu		
	130					135					140						
Phe	Lys	Asn	Gly	Glu	Lys	Arg	Glu	Ser	Ile	Met	Gly	Thr	Met	Pro	Lys		
145					150					155					160		
Glu	Phe	Tyr	Ile	Ser	Ala	Ile	Glu	Arg	Val	Leu	Asn	Ser					
				165					170								

<210> 62

<211> 193

<212> PRT

<213> Arabidopsis thaliana

<400> 62

Met	Ala	Ser	Leu	Leu	Asp	Ser	Val	Thr	Val	Thr	Arg	Val	Phe	Ser	Leu		
1				5					10					15			
Pro	Ile	Ala	Ala	Ser	Val	Ser	Ser	Ser	Ser	Ala	Ala	Pro	Ser	Val	Ser		
			20					25					30				
Arg	Arg	Arg	Ile	Ser	Pro	Ala	Arg	Phe	Leu	Glu	Phe	Arg	Gly	Leu	Lys		
		35					40					45					
Ser	Ser	Arg	Ser	Leu	Val	Thr	Gln	Ser	Ala	Ser	Leu	Gly	Ala	Asn	Arg		
	50					55					60						
Arg	Thr	Arg	Ile	Ala	Arg	Gly	Gly	Arg	Ile	Ala	Cys	Glu	Ala	Gln	Asp		
65					70					75					80		
Thr	Thr	Ala	Ala	Ala	Val	Glu	Val	Pro	Asn	Leu	Ser	Asp	Ser	Glu	Trp		
				85				90						95			
Gln	Thr	Lys	Val	Leu	Glu	Ser	Asp	Val	Pro	Val	Leu	Val	Glu	Phe	Trp		
			100					105					110				
Ala	Pro	Trp	Cys	Gly	Pro	Cys	Arg	Met	Ile	His	Pro	Ile	Val	Asp	Gln		
	115						120					125					
Leu	Ala	Lys	Asp	Phe	Ala	Gly	Lys	Phe	Lys	Phe	Tyr	Lys	Ile	Asn	Thr		
	130					135					140						
Asp	Glu	Ser	Pro	Asn	Thr	Pro	Asn	Arg	Tyr	Gly	Ile	Arg	Ser	Val	Pro		
145					150					155					160		

Thr Val Ile Ile Phe Lys Gly Gly Glu Lys Lys Asp Ser Ile Ile Gly
 165 170 175
 Ala Val Pro Arg Glu Thr Leu Glu Lys Thr Ile Glu Arg Phe Leu Val
 180 185 190
 Glu

<210> 63.
 <211> 177
 <212> PRT
 <213> Brassica napus

<400> 63
 Met Ala Ala Phe Thr Cys Thr Ser Ser Pro Pro Ile Ser Leu Arg Ser
 1 5 10 15
 Glu Met Met Ile Ala Ser Ser Lys Thr Val Ser Leu Ser Thr Arg Gln
 20 25 30
 Met Phe Ser Val Gly Gly Leu Arg Thr Arg Val Ser Leu Ser Ser Val
 35 40 45
 Ser Lys Asn Ser Arg Ala Ser Arg Leu Arg Arg Gly Gly Ile Ile Cys
 50 55 60
 Glu Ala Gln Asp Thr Ala Thr Gly Ile Pro Met Val Asn Asp Ser Thr
 65 70 75 80
 Trp Glu Ser Leu Val Leu Lys Ala Asp Glu Pro Val Val Val Asp Phe
 85 90 95
 Trp Ala Pro Trp Cys Gly Pro Cys Lys Met Ile Asp Pro Ile Val Asn
 100 105 110
 Glu Leu Ala Gln Gln Tyr Thr Gly Lys Ile Lys Phe Phe Lys Leu Asn
 115 120 125
 Thr Asp Asp Ser Pro Ala Thr Pro Gly Lys Tyr Gly Val Arg Ser Ile
 130 135 140
 Pro Thr Ile Met Ile Phe Val Lys Gly Glu Lys Lys Asp Thr Ile Ile
 145 150 155 160
 Gly Ala Val Pro Lys Thr Thr Leu Ala Thr Ser Ile Asp Lys Phe Leu
 165 170 175
 Gln

<210> 64
 <211> 140
 <212> PRT
 <213> Chlamydomonas reinhardtii

<400> 64
 Met Ala Leu Val Ala Arg Arg Ala Ala Val Pro Ser Ala Arg Ser Ser
 1 5 10 15
 Ala Arg Pro Ala Phe Ala Arg Ala Ala Pro Arg Arg Ser Val Val Val
 20 25 30
 Arg Ala Glu Ala Gly Ala Val Asn Asp Asp Thr Phe Lys Asn Val Val
 35 40 45
 Leu Glu Ser Ser Val Pro Val Leu Val Asp Phe Trp Ala Pro Trp Cys
 50 55 60
 Gly Pro Cys Arg Ile Ile Ala Pro Val Val Asp Glu Ile Ala Gly Glu
 65 70 75 80
 Tyr Lys Asp Lys Leu Lys Cys Val Lys Leu Asn Thr Asp Glu Ser Pro
 85 90 95
 Asn Val Ala Ser Glu Tyr Gly Ile Arg Ser Ile Pro Thr Ile Met Val
 100 105 110
 Phe Lys Gly Gly Lys Lys Cys Glu Thr Ile Ile Gly Ala Val Pro Lys
 115 120 125
 Ala Thr Ile Val Gln Thr Val Glu Lys Tyr Leu Asn
 130 135 140

<210> 65

<211> 167
 <212> PRT
 <213> Zea mays

<400> 65
 Met Ala Met Glu Thr Cys Phe Arg Ala Trp Ala Leu His Ala Pro Ala
 1 5 10 15
 Gly Ser Lys Asp Arg Leu Leu Val Gly Asn Leu Val Leu Pro Ser Lys
 20 25 30
 Arg Ala Leu Ala Pro Leu Ser Val Gly Arg Val Ala Thr Arg Arg Pro
 35 40 45
 Arg His Val Cys Gln Ser Lys Asn Ala Val Asp Glu Val Val Val Ala
 50 55 60
 Asp Glu Lys Asn Trp Asp Gly Leu Val Met Ala Cys Glu Thr Pro Val
 65 70 75 80
 Leu Val Glu Phe Trp Ala Pro Trp Cys Gly Pro Cys Arg Met Ile Ala
 85 90 95
 Pro Val Ile Asp Glu Leu Ala Lys Asp Tyr Ala Gly Lys Ile Thr Cys
 100 105 110
 Cys Lys Val Asn Thr Asp Asp Ser Pro Asn Val Ala Ser Thr Tyr Gly
 115 120 125
 Ile Arg Ser Ile Pro Thr Val Leu Ile Phe Lys Gly Gly Glu Lys Lys
 130 135 140
 Glu Ser Val Ile Gly Ala Val Pro Lys Ser Thr Leu Thr Thr Leu Ile
 145 150 155 160
 Asp Lys Tyr Ile Gly Ser Ser
 165

<210> 66
 <211> 172
 <212> PRT
 <213> Oryza sativa

<400> 66
 Met Ala Leu Glu Thr Cys Phe Arg Ala Trp Ala Thr Leu His Ala Pro
 1 5 10 15
 Gln Pro Pro Ser Ser Gly Gly Ser Arg Asp Arg Leu Leu Leu Ser Gly
 20 25 30
 Ala Gly Ser Ser Gln Ser Lys Pro Arg Leu Ser Val Ala Ser Pro Ser
 35 40 45
 Pro Leu Arg Pro Ala Ser Arg Phe Ala Cys Gln Cys Ser Asn Val Val
 50 55 60
 Asp Glu Val Val Val Ala Asp Glu Lys Asn Trp Asp Ser Met Val Leu
 65 70 75 80
 Gly Ser Glu Ala Pro Val Leu Val Glu Phe Trp Ala Pro Trp Cys Gly
 85 90 95
 Pro Cys Arg Met Ile Ala Pro Val Ile Asp Glu Leu Ala Lys Glu Tyr
 100 105 110
 Val Gly Lys Ile Lys Cys Cys Lys Val Asn Thr Asp Asp Ser Pro Asn
 115 120 125
 Ile Ala Thr Asn Tyr Gly Ile Arg Ser Ile Pro Thr Val Leu Met Phe
 130 135 140
 Lys Asn Gly Glu Lys Lys Glu Ser Val Ile Gly Ala Val Pro Lys Thr
 145 150 155 160
 Thr Leu Ala Thr Ile Ile Asp Lys Tyr Val Ser Ser
 165 170

<210> 67
 <211> 172
 <212> PRT
 <213> Pisum sativum

<400> 67
 Met Ala Leu Glu Ser Leu Phe Lys Ser Ile His Thr Lys Thr Ser Leu
 1 5 10 15

Ser Ser Ser Ile Val Phe Ile Phe Lys Gly Lys Ala Cys Leu Leu Thr
 20 25 30
 Ser Lys Ser Arg Ile Gln Glu Ser Phe Ala Glu Leu Asn Ser Phe Thr
 35 40 45
 Ser Leu Val Leu Leu Ile Glu Asn His Val Leu Leu His Ala Arg Glu
 50 55 60
 Ala Val Asn Glu Val Gln Val Val Asn Asp Ser Ser Trp Asp Glu Leu
 65 70 75 80
 Val Ile Gly Ser Glu Thr Pro Val Leu Val Asp Phe Trp Ala Pro Trp
 85 90 95
 Cys Gly Pro Cys Arg Met Ile Ala Pro Ile Ile Asp Glu Leu Ala Lys
 100 105 110
 Glu Tyr Ala Gly Lys Ile Lys Cys Tyr Lys Leu Asn Thr Asp Glu Ser
 115 120 125
 Pro Asn Thr Ala Thr Lys Tyr Gly Ile Arg Ser Ile Pro Thr Val Leu
 130 135 140
 Phe Phe Lys Asn Gly Glu Arg Lys Asp Ser Val Ile Gly Ala Val Pro
 145 150 155 160
 Lys Ala Thr Leu Ser Glu Lys Val Glu Lys Tyr Ile
 165 170

<210> 68
 <211> 181
 <212> PRT
 <213> Spinacia oleracea

<400> 68
 Met Ala Ile Glu Asn Cys Leu Gln Leu Ser Thr Ser Ala Ser Val Gly
 1 5 10 15
 Thr Val Ala Val Lys Ser His Val His His Leu Gln Pro Ser Ser Lys
 20 25 30
 Val Asn Val Pro Thr Phe Arg Gly Leu Lys Arg Ser Phe Pro Ala Leu
 35 40 45
 Ser Ser Ser Val Ser Ser Ser Pro Arg Gln Phe Arg Tyr Ser Ser
 50 55 60
 Val Val Cys Lys Ala Ser Glu Ala Val Lys Glu Val Gln Asp Val Asn
 65 70 75 80
 Asp Ser Ser Trp Lys Glu Phe Val Leu Glu Ser Glu Val Pro Val Met
 85 90 95
 Val Asp Phe Trp Ala Pro Trp Cys Gly Pro Cys Lys Leu Ile Ala Pro
 100 105 110
 Val Ile Asp Glu Leu Ala Lys Glu Tyr Ser Gly Lys Ile Ala Val Tyr
 115 120 125
 Lys Leu Asn Thr Asp Glu Ala Pro Gly Ile Ala Thr Gln Tyr Asn Ile
 130 135 140
 Arg Ser Ile Pro Thr Val Leu Phe Phe Lys Asn Gly Glu Arg Lys Glu
 145 150 155 160
 Ser Ile Ile Gly Ala Val Pro Lys Ser Thr Leu Thr Asp Ser Ile Glu
 165 170 175
 Lys Tyr Leu Ser Pro
 180

<210> 69
 <211> 175
 <212> PRT
 <213> Triticum aestivum

<400> 69
 Met Ala Leu Glu Thr Cys Leu Arg Gly Trp Ala Leu Tyr Ala Pro Gln
 1 5 10 15
 Ala Gly Ile Arg Glu Arg Leu Ser Ser Gly Ser Tyr Ala Pro Ser Arg
 20 25 30
 Pro Arg Thr Ala Ala Pro Ala Val Ser Pro Ser Pro Tyr Lys Ser
 35 40 45
 Ala Leu Val Ala Ala Arg Arg Pro Ser Arg Phe Val Cys Lys Cys Lys

50	55	60
Asn Val Val Asp Glu Val Ile Val Ala Asp Glu Lys Asn Trp Asp Asn		
65	70	75
Met Val Ile Ala Cys Glu Ser Pro Val Leu Val Glu Phe Trp Ala Pro		80
	85	90
Trp Cys Gly Pro Cys Arg Met Ile Ala Pro Val Ile Asp Glu Leu Ala		95
	100	105
Lys Asp Tyr Val Gly Lys Ile Lys Cys Cys Lys Val Asn Thr Asp Asp		110
	115	120
Cys Pro Asn Ile Ala Ser Thr Tyr Gly Ile Arg Ser Ile Pro Thr Val		125
	130	135
Leu Met Phe Lys Asp Gly Glu Lys Lys Glu Ser Val Ile Gly Ala Val		140
145	150	155
Pro Lys Thr Thr Leu Cys Thr Ile Ile Asp Lys Tyr Ile Gly Ser		160
	165	170
		175

<210> 70
 <211> 106
 <212> PRT
 <213> Anacystis nidulans

<400> 70
Ser Val Ala Ala Ala Val Thr Asp Ala Thr Phe Lys Gln Glu Val Leu
1 5 10 15
Glu Ser Ser Ile Pro Val Leu Val Asp Phe Trp Ala Pro Trp Cys Gly
20 25 30
Pro Cys Arg Met Val Ala Pro Val Val Asp Glu Ile Ala Gln Gln Tyr
35 40 45
Ser Asp Gln Val Lys Val Val Lys Val Asn Thr Asp Glu Asn Pro Ser
50 55 60
Val Ala Ser Gln Tyr Gly Ile Arg Ser Ile Pro Thr Leu Met Ile Phe
65 70 75 80
Lys Asp Gly Gln Arg Val Asp Thr Val Val Gly Ala Val Pro Lys Thr
85 90 95
Thr Leu Ala Asn Thr Leu Asp Lys His Leu
100 105

<210> 71
 <211> 107
 <212> PRT
 <213> Cyanidium caldarium

<400> 71
Met Pro Ser Pro Ile Gln Val Thr Asp Phe Ser Phe Glu Lys Glu Val
1 5 10 15
Val Asn Ser Glu Lys Leu Val Leu Val Asp Phe Trp Ala Pro Trp Cys
20 25 30
Gly Pro Cys Arg Met Ile Ser Pro Val Ile Asp Glu Leu Ala Gln Glu
35 40 45
Tyr Val Glu Gln Val Lys Ile Val Lys Ile Asn Thr Asp Glu Asn Pro
50 55 60
Ser Ile Ser Ala Glu Tyr Gly Ile Arg Ser Ile Pro Thr Leu Met Leu
65 70 75 80
Phe Lys Asp Gly Lys Arg Val Asp Thr Val Ile Gly Ala Val Pro Lys
85 90 95
Ser Thr Leu Thr Asn Ala Leu Lys Lys Tyr Leu
100 105

<210> 72
 <211> 102
 <212> PRT
 <213> Cyanidioschyzon merolae

<400> 72

```

Met Leu His Ile Asp Glu Leu Thr Phe Glu Asn Glu Val Leu Gln Ser
 1          5          10          15
Glu Lys Leu Val Leu Val Asp Phe Trp Ala Pro Trp Cys Gly Pro Cys
          20          25          30
Arg Met Ile Gly Pro Ile Leu Glu Glu Ile Ala Lys Glu Phe Asn Leu
 35          40          45
Lys Val Val Gln Val Asn Thr Asp Glu Asn Pro Asn Leu Ala Thr Phe
 50          55          60
Tyr Gly Ile Arg Ser Ile Pro Thr Leu Met Leu Phe Lys Lys Gly Gln
65          70          75          80
Arg Val Asp Thr Val Ile Gly Ala Val Pro Lys Ser Ile Leu Ile His
          85          90          95
Thr Ile Asn Lys Tyr Leu
          100

```

<210> 73
 <211> 109
 <212> PRT
 <213> *Griffithsia pacifica*

```

<400> 73
Met Ser Ile Ser Gln Val Ile Asp Thr Ser Phe His Glu Glu Val Ile
 1          5          10          15
Asn Ser Arg Gln Pro Val Leu Val Asp Phe Trp Ala Pro Trp Cys Gly
          20          25          30
Pro Cys Arg Met Ile Ala Ser Thr Ile Asp Glu Ile Ala His Asp Tyr
          35          40          45
Lys Asp Lys Leu Lys Val Val Lys Val Asn Thr Asp Gln Asn Pro Thr
 50          55          60
Ile Ala Thr Glu Tyr Gly Ile Arg Ser Ile Pro Thr Val Met Ile Phe
65          70          75          80
Ile Asn Gly Lys Lys Val Asp Thr Val Val Gly Ala Val Pro Lys Leu
          85          90          95
Thr Leu Leu Asn Thr Leu Gln Lys His Leu Lys Ser Thr
          100          105

```

<210> 74
 <211> 107
 <212> PRT
 <213> *Porphyra yezoensis*

```

<400> 74
Met Ser Val Ser Gln Val Thr Asp Ala Ser Phe Lys Gln Glu Val Ile
 1          5          10          15
Asn Asn Asn Leu Pro Val Leu Val Asp Phe Trp Ala Pro Trp Cys Gly
          20          25          30
Pro Cys Arg Met Val Ser Pro Val Val Asp Glu Ile Ala Glu Glu Tyr
          35          40          45
Glu Ser Ser Ile Lys Val Val Lys Ile Asn Thr Asp Asp Asn Pro Thr
 50          55          60
Ile Ala Ala Glu Tyr Gly Ile Arg Ser Ile Pro Thr Leu Met Ile Phe
65          70          75          80
Lys Ala Gly Glu Arg Val Asp Thr Val Ile Gly Ala Val Pro Lys Ser
          85          90          95
Thr Leu Ala Ser Thr Leu Asn Lys Tyr Ile Ser
          100          105

```

<210> 75
 <211> 107
 <212> PRT
 <213> *Porphyra purpurea*

```

<400> 75
Met Ser Val Ser Gln Val Thr Asp Ala Ser Phe Lys Gln Glu Val Ile

```



```

1      5      10      15
Asn Asn Asp Leu Pro Val Leu Val Asp Phe Trp Ala Pro Trp Cys Gly
20      25      30
Pro Cys Arg Met Val Ser Pro Val Val Asp Ala Ile Ala Glu Glu Tyr
35      40      45
Glu Ser Ser Ile Lys Val Val Lys Ile Asn Thr Asp Asp Asn Pro Thr
50      55      60
Ile Ala Ala Glu Tyr Gly Ile Arg Ser Ile Pro Thr Leu Met Ile Phe
65      70      75      80
Lys Ser Gly Glu Arg Val Asp Thr Val Ile Gly Ala Val Pro Lys Ser
85      90      95
Thr Leu Glu Ser Thr Leu Asn Lys Tyr Ile Ser
100      105

```

<210> 76
 <211> 114
 <212> PRT
 <213> Arabidopsis thaliana

```

<400> 76
Met Ala Ser Glu Glu Gly Gln Val Ile Ala Cys His Thr Val Glu Thr
1      5      10      15
Trp Asn Glu Gln Leu Gln Lys Ala Asn Glu Ser Lys Thr Leu Val Val
20      25      30
Val Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro
35      40      45
Phe Phe Ala Asp Leu Ala Lys Lys Leu Pro Asn Val Leu Phe Leu Lys
50      55      60
Val Asp Thr Asp Glu Leu Lys Ser Val Ala Ser Asp Trp Ala Ile Gln
65      70      75      80
Ala Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Leu Asp Lys
85      90      95
Val Val Gly Ala Lys Lys Asp Glu Leu Gln Ser Thr Ile Ala Lys His
100      105      110
Leu Ala

```

<210> 77
 <211> 110
 <212> PRT
 <213> Anabaena

```

<400> 77
Ser Lys Gly Val Ile Thr Ile Thr Asp Ala Glu Phe Glu Ser Glu Val
1      5      10      15
Leu Lys Ala Glu Gln Pro Val Leu Val Tyr Phe Trp Ala Ser Trp Cys
20      25      30
Gly Pro Cys Gln Leu Met Ser Pro Leu Ile Asn Leu Ala Asn Thr
35      40      45
Tyr Ser Asp Arg Leu Lys Val Val Lys Leu Glu Ile Asp Pro Asn Pro
50      55      60
Thr Thr Val Lys Lys Tyr Lys Val Glu Gly Val Pro Ala Leu Arg Leu
65      70      75      80
Val Lys Gly Glu Gln Ile Leu Asp Ser Thr Glu Gly Val Ile Ser Lys
85      90      95
Asp Lys Leu Leu Ser Phe Leu Asp Thr His Leu Asn Asn Asn
100      105      110

```

<210> 78
 <211> 123
 <212> PRT
 <213> Brassica napus

<400> 78

```

Met Ala Ala Thr Ala Glu Val Ile Pro Ala Gly Glu Val Ile Ala Cys
 1          5          10          15
His Thr Val Glu Asp Trp Asn Asn Lys Leu Lys Ala Ala Lys Glu Ser
          20          25          30
Asn Lys Leu Ile Val Ile Asp Phe Thr Ala Val Trp Cys Pro Fro Cys
          35          40          45
Arg Phe Ile Ala Pro Ile Phe Val Glu Leu Ala Lys Lys His Leu Asp
          50          55          60
Val Val Phe Phe Lys Val Asp Val Asp Glu Leu Ala Thr Val Ala Gln
65          70          75          80
Glu Phe Asp Val Gln Ala Met Pro Thr Phe Val Tyr Met Lys Gly Glu
          85          90          95
Glu Lys Leu Asp Lys Val Val Gly Ala Ala Lys Glu Glu Ile Glu Ala
          100          105          110
Lys Leu Leu Lys His Ser Gln Val Ala Ala Ala
          115          120

```

<210> 79
 <211> 126
 <212> PRT
 <213> *Nicotiana tabacum*

```

<400> 79
Met Ala Ala Asn Asp Ala Thr Ser Ser Glu Glu Gly Gln Val Phe Gly
 1          5          10          15
Cys His Lys Val Glu Glu Trp Asn Glu Tyr Phe Lys Lys Gly Val Glu
          20          25          30
Thr Lys Lys Leu Val Val Val Asp Phe Thr Ala Ser Trp Cys Gly Pro
          35          40          45
Cys Arg Phe Ile Ala Pro Ile Leu Ala Asp Ile Ala Lys Lys Met Pro
          50          55          60
His Val Ile Phe Leu Lys Val Asp Val Asp Glu Leu Lys Thr Val Ser
65          70          75          80
Ala Glu Trp Ser Val Glu Ala Met Pro Thr Phe Val Phe Ile Lys Asp
          85          90          95
Gly Lys Glu Val Asp Arg Val Val Gly Ala Lys Lys Glu Glu Leu Gln
          100          105          110
Gln Thr Ile Val Lys His Ala Ala Pro Ala Thr Val Thr Ala
          115          120          125

```

<210> 80
 <211> 133
 <212> PRT
 <213> *Arabidopsis thaliana*

```

<400> 80
Met Gly Gly Ala Leu Ser Thr Val Phe Gly Ser Gly Glu Asp Ala Thr
 1          5          10          15
Ala Ala Gly Thr Glu Ser Glu Pro Ser Arg Val Leu Lys Phe Ser Ser
          20          25          30
Ser Ala Arg Trp Gln Leu His Phe Asn Glu Ile Lys Glu Ser Asn Lys
          35          40          45
Leu Leu Val Val Asp Phe Ser Ala Ser Trp Cys Gly Pro Cys Arg Met
          50          55          60
Ile Glu Pro Ala Ile His Ala Met Ala Asp Lys Phe Asn Asp Val Asp
65          70          75          80
Phe Val Lys Leu Asp Val Asp Glu Leu Pro Asp Val Ala Lys Glu Phe
          85          90          95
Asn Val Thr Ala Met Pro Thr Phe Val Leu Val Lys Arg Gly Lys Glu
          100          105          110
Ile Glu Arg Ile Ile Gly Ala Lys Lys Asp Glu Leu Glu Lys Lys Val
          115          120          125
Ser Lys Leu Arg Ala
          130

```

<210> 81
 <211> 119
 <212> PRT
 <213> Brassica napus

<400> 81
 Met Ala Ala Glu Glu Gly Gln Val Ile Gly Cys His Glu Ile Asp Val
 1 5 10 15
 Trp Ala Val Gln Leu Asp Thr Ala Lys Gln Ser Asn Lys Leu Ile Val
 20 25 30
 Ile Asp Phe Thr Ala Ser Trp Cys Pro Pro Cys Arg Met Ile Ala Pro
 35 40 45
 Val Phe Ala Asp Leu Ala Lys Lys Phe Met Ser Ser Ala Ile Phe Phe
 50 55 60
 Lys Val Asp Val Asp Glu Leu Gln Asn Val Ala Gln Glu Phe Gly Val
 65 70 75 80
 Glu Ala Met Pro Thr Phe Val Leu Ile Lys Asp Gly Asn Val Val Asp
 85 90 95
 Lys Val Val Gly Ala Arg Lys Glu Asp Leu His Ala Thr Ile Ala Lys
 100 105 110
 His Thr Gly Val Ala Thr Ala
 115

<210> 82
 <211> 118
 <212> PRT
 <213> Nicotiana tabacum

<400> 82
 Met Ala Glu Glu Gly Gln Val Ile Gly Val His Thr Val Asp Ala Trp
 1 5 10 15
 Asn Glu His Leu Gln Lys Gly Ile Asp Asp Lys Lys Leu Ile Val Val
 20 25 30
 Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Lys Phe Ile Ala Ser Phe
 35 40 45
 Tyr Ala Glu Leu Ala Lys Lys Met Pro Thr Val Thr Phe Leu Lys Val
 50 55 60
 Asp Val Asp Glu Leu Lys Ser Val Ala Thr Asp Trp Ala Val Glu Ala
 65 70 75 80
 Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Val Asp Lys Val
 85 90 95
 Val Gly Ala Lys Lys Asp Glu Leu Gln Gln Thr Ile Ala Lys His Ile
 100 105 110
 Ser Ser Thr Ser Thr Ala
 115

<210> 83
 <211> 118
 <212> PRT
 <213> Arabidopsis thaliana

<400> 83
 Met Ala Ala Glu Gly Glu Val Ile Ala Cys His Thr Val Glu Asp Trp
 1 5 10 15
 Thr Glu Lys Leu Lys Ala Ala Asn Glu Ser Lys Lys Leu Ile Val Ile
 20 25 30
 Asp Phe Thr Ala Thr Trp Cys Pro Pro Cys Arg Phe Ile Ala Pro Val
 35 40 45
 Phe Ala Asp Leu Ala Lys Lys His Leu Asp Val Val Phe Phe Lys Val
 50 55 60
 Asp Val Asp Glu Leu Asn Thr Val Ala Glu Glu Phe Lys Val Gln Ala
 65 70 75 80
 Met Pro Thr Phe Ile Phe Met Lys Glu Gly Glu Ile Lys Glu Thr Val
 85 90 95
 Val Gly Ala Ala Lys Glu Glu Ile Ile Ala Asn Leu Glu Lys His Lys

Thr Val Val Ala Ala Ala 100 105 110
115

<210> 84
<211> 125
<212> PRT
<213> Arabidopsis thaliana

<400> 84
Met Ala Ala Glu Glu Gly Gln Val Ile Gly Cys His Thr Asn Asp Val
1 5 10 15
Trp Thr Val Gln Leu Asp Lys Ala Lys Glu Ser Asn Lys Leu Ile Val
20 25 30
Ile Asp Phe Thr Ala Ser Trp Cys Pro Pro Cys Arg Met Ile Ala Pro
35 40 45
Ile Phe Asn Asp Leu Ala Lys Lys Phe Met Ser Ser Ala Ile Phe Phe
50 55 60
Lys Val Asp Val Asp Glu Leu Gln Ser Val Ala Lys Glu Phe Gly Val
65 70 75 80
Glu Ala Met Pro Thr Phe Val Phe Ile Lys Ala Gly Glu Val Val Asp
85 90 95
Lys Leu Val Gly Ala Asn Lys Glu Asp Leu Gln Ala Lys Ile Val Lys
100 105 110
His Thr Gly Val Thr Thr Val Val Asn Gln Phe Glu Ala
115 120 125

<210> 85
<211> 118
<212> PRT
<213> Arabidopsis thaliana

<400> 85
Met Ala Gly Glu Gly Glu Val Ile Ala Cys His Thr Leu Glu Val Trp
1 5 10 15
Asn Glu Lys Val Lys Asp Ala Asn Glu Ser Lys Lys Leu Ile Val Ile
20 25 30
Asp Phe Thr Ala Ser Trp Cys Pro Pro Cys Arg Phe Ile Ala Pro Val
35 40 45
Phe Ala Glu Met Ala Lys Lys Phe Thr Asn Val Val Phe Phe Lys Ile
50 55 60
Asp Val Asp Glu Leu Gln Ala Val Ala Gln Glu Phe Lys Val Glu Ala
65 70 75 80
Met Pro Thr Phe Val Phe Met Lys Glu Gly Asn Ile Ile Asp Arg Val
85 90 95
Val Gly Ala Ala Lys Asp Glu Ile Asn Glu Lys Leu Met Lys His Gly
100 105 110
Gly Leu Val Ala Ser Ala
115

<210> 86
<211> 123
<212> PRT
<213> Brassica rapa

<400> 86
Met Ala Ala Thr Ala Glu Leu Ile Pro Ala Gly Glu Val Ile Ala Cys
1 5 10 15
His Thr Val Glu Asp Trp Asn Asn Lys Leu Lys Ala Ala Lys Glu Ser
20 25 30
Asn Lys Leu Ile Val Ile Asp Phe Thr Ala Val Trp Cys Pro Pro Cys
35 40 45
Arg Phe Ile Ala Pro Ile Phe Val Glu Leu Ala Lys Lys His Leu Asp
50 55 60

Val Val Phe Phe Lys Val Asp Val Asp Glu Leu Ala Thr Val Ala Lys
 65 70 75 80
 Glu Phe Asp Val Gln Ala Met Pro Thr Phe Val Tyr Met Lys Gly Glu
 85 90 95
 Glu Lys Leu Asp Lys Val Val Gly Ala Ala Lys Glu Glu Ile Glu Ala
 100 105 110
 Lys Leu Leu Lys His Ser Gln Val Ala Ala Ala
 115 120

<210> 87
 <211> 112
 <212> PRT
 <213> Chlamydomonas reinhardtii

<400> 87
 Gly Gly Ser Val Ile Val Ile Asp Ser Lys Ala Ala Trp Asp Ala Gln
 1 5 10 15
 Leu Ala Lys Gly Lys Glu Glu His Lys Pro Ile Val Val Asp Phe Thr
 20 25 30
 Ala Thr Trp Cys Gly Pro Cys Lys Met Ile Ala Pro Leu Phe Glu Thr
 35 40 45
 Leu Ser Asn Asp Tyr Ala Gly Lys Val Ile Phe Leu Lys Val Asp Val
 50 55 60
 Asp Ala Val Ala Ala Val Ala Glu Ala Ala Gly Ile Thr Ala Met Pro
 65 70 75 80
 Thr Phe His Val Tyr Lys Asp Gly Val Lys Ala Asp Asp Leu Val Gly
 85 90 95
 Ala Ser Gln Asp Lys Leu Lys Ala Leu Val Ala Lys His Ala Ala Ala
 100 105 110

<210> 88
 <211> 116
 <212> PRT
 <213> Fagopyrum esculentum

<400> 88
 Met Ala Glu Glu Ala Gln Val Ile Ala Cys His Thr Val Gln Glu Trp
 1 5 10 15
 Asn Glu Lys Phe Gln Lys Ala Lys Asp Ser Gly Lys Leu Ile Val Ile
 20 25 30
 Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Val Ile Thr Pro Tyr
 35 40 45
 Val Ser Glu Leu Ala Lys Lys Phe Pro His Val Ala Phe Phe Lys Val
 50 55 60
 Asp Val Asp Asp Leu Lys Asp Val Ala Glu Glu Tyr Lys Val Glu Ala
 65 70 75 80
 Met Pro Ser Phe Val Ile Leu Lys Glu Gly Gln Glu Val Glu Arg Ile
 85 90 95
 Val Gly Ala Arg Lys Asp Glu Leu Leu His Lys Ile Ala Val His Ala
 100 105 110
 Pro Ile Thr Ala
 115

<210> 89
 <211> 122
 <212> PRT
 <213> Oryza sativa

<400> 89
 Met Ala Ala Glu Glu Gly Val Val Ile Ala Cys His Asn Lys Asp Glu
 1 5 10 15
 Phe Asp Ala Gln Met Thr Lys Ala Lys Glu Ala Gly Lys Val Val Ile
 20 25 30
 Ile Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro

```

      35      40      45
Val Phe Ala Glu Tyr Ala Lys Lys Phe Pro Gly Ala Val Phe Leu Lys
    50      55      60
Val Asp Val Asp Glu Leu Lys Glu Val Ala Glu Lys Tyr Asn Val Glu
    65      70      75      80
Ala Met Pro Thr Phe Leu Phe Ile Lys Asp Gly Ala Glu Ala Asp Lys
      85      90      95
Val Val Gly Ala Arg Lys Asp Asp Leu Gln Asn Thr Ile Val Lys His
      100      105      110
Val Gly Ala Thr Ala Ala Ser Ala Ser Ala
      115      120

```

<210> 90
 <211> 125
 <212> PRT
 <213> Picea mariana

```

<400> 90
Met Ala Glu Gly Asn Val Phe Ala Cys His Ser Thr Glu Gly Trp Arg
  1      5      10      15
Ser Lys Leu Gln Glu Ala Ile Asp Thr Lys Arg Leu Val Ala Val Asp
      20      25      30
Phe Thr Ala Thr Trp Cys Gly Pro Cys Arg Val Ile Gly Pro Val Phe
      35      40      45
Val Glu Leu Ser Lys Lys Phe Pro Glu Ile Phe Phe Leu Lys Val Asp
      50      55      60
Val Asp Glu Leu Arg Asp Val Ala Gln Glu Trp Asp Val Glu Ala Met
      65      70      75      80
Pro Thr Phe Ile Phe Ile Lys Asp Gly Lys Ala Val Asp Lys Val Val
      85      90      95
Gly Ala Lys Lys Asp Asp Leu Glu Arg Lys Val Ala Ala Leu Ala Ala
      100      105      110
Ala Ala Thr Thr Thr Glu Ala Thr Leu Pro Ala Gln Ala
      115      120      125

```

<210> 91
 <211> 118
 <212> PRT
 <213> Ricinus communis

```

<400> 91
Met Ala Ala Glu Glu Gly Gln Val Ile Gly Cys His Thr Val Glu Ala
  1      5      10      15
Trp Asn Glu Gln Leu Gln Lys Gly Asn Asp Thr Lys Gly Leu Ile Val
      20      25      30
Val Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro
      35      40      45
Phe Leu Ala Glu Leu Ala Lys Lys Leu Pro Asn Val Thr Phe Leu Lys
      50      55      60
Val Asp Val Asp Glu Leu Lys Thr Val Ala His Glu Trp Ala Val Glu
      65      70      75      80
Ser Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Met Asp Lys
      85      90      95
Val Val Gly Ala Lys Lys Asp Glu Leu Gln Gln Thr Ile Ala Lys His
      100      105      110
Met Ala Thr Ala Ser Thr
      115

```

<210> 92
 <211> 126
 <212> PRT
 <213> triticum aestivum

<400> 92

Ala Ala Ser Ala Ala Thr Ala Thr Ala Thr Ala Ala Ala Val Gly Ala
 1 5 10 15
 Gly Glu Val Ile Ser Val His Ser Leu Glu Gln Trp Thr Met Gln Ile
 20 25 30
 Glu Glu Ala Asn Ala Ala Lys Lys Leu Val Val Ile Asp Phe Thr Ala
 35 40 45
 Ser Trp Cys Gly Pro Cys Arg Ile Met Ala Pro Ile Phe Ala Asp Leu
 50 55 60
 Ala Lys Lys Phe Pro Ala Ala Val Phe Leu Lys Val Asp Val Asp Glu
 65 70 75 80
 Leu Lys Pro Ile Ala Glu Gln Phe Ser Val Glu Ala Met Pro Thr Phe
 85 90 95
 Leu Phe Met Lys Glu Gly Asp Val Lys Asp Arg Val Val Gly Ala Ile
 100 105 110
 Lys Glu Glu Leu Thr Thr Lys Val Gly Leu His Ala Ala Gln
 115 120 125

<210> 93

<211> 109

<212> PRT

<213> *Aspergillus nidulans*

<400> 93

Gly Ala Ser Glu His Val Pro Pro Ile Thr Ser Lys Ala Glu Phe Gln
 1 5 10 15
 Glu Lys Val Leu Asn Ala Lys Gly Phe Val Val Val Asp Cys Phe Ala
 20 25 30
 Thr Trp Cys Gly Pro Cys Lys Ala Ile Ala Pro Thr Val Glu Lys Phe
 35 40 45
 Ala Gln Thr Tyr Thr Asp Ala Ser Phe Tyr Gln Ile Asp Val Asp Glu
 50 55 60
 Leu Ser Glu Val Ala Ala Glu Leu Gly Ile Arg Ala Met Pro Thr Phe
 65 70 75 80
 Leu Leu Phe Lys Asp Gly Gln Lys Val Ser Asp Val Val Gly Ala Asn
 85 90 95
 Pro Gly Ala Leu Glu Ala Gly Ile Lys Ala Leu Leu Ala
 100 105

<210> 94

<211> 105

<212> PRT

<213> *Alicyclobacillus*

<400> 94

Ala Thr Met Thr Leu Thr Asp Ala Asn Phe Gln Gln Ala Ile Gln Gly
 1 5 10 15
 Asp Lys Pro Val Leu Val Asp Phe Trp Ala Ala Trp Cys Gly Pro Cys
 20 25 30
 Arg Met Met Ala Pro Val Leu Glu Glu Phe Ala Glu Ala His Ala Asp
 35 40 45
 Lys Val Thr Val Ala Lys Leu Asn Val Asp Glu Asn Pro Glu Thr Thr
 50 55 60
 Ser Gln Phe Gly Ile Met Ser Ile Pro Thr Leu Ile Leu Phe Lys Gly
 65 70 75 80
 Gly Arg Pro Val Lys Gln Leu Ile Gly Tyr Gln Pro Lys Glu Gln Leu
 85 90 95
 Glu Ala Gln Leu Ala Asp Val Leu Gln
 100 105

<210> 95

<211> 91

<212> PRT

<213> *Archaeoglobus fulgidus*

<400> 95

```

Met Val Met Met Lys Leu Phe Thr Ser Pro Thr Cys Pro Tyr Cys Pro
 1          5          10          15
Lys Ala Glu Lys Val Val Ser Lys Val Ala Lys Glu Glu Gly Val Leu
          20          25          30
Ala Ile Asn Leu Pro Val Asn Thr Asp Glu Gly Leu Lys Glu Ala Leu
          35          40          45
Lys Phe Gly Ile Arg Gly Val Pro Ala Leu Val Ile Asn Asp Lys Tyr
          50          55          60
Leu Ile Leu Gly Val Pro Asp Glu Gly Glu Leu Arg Gln Leu Ile Arg
65          70          75          80
Lys Leu Lys Gly Gly Glu Glu Tyr Gly Ala Ser
          85          90

```

<210> 96

<211> 103

<212> PRT

<213> *Bacillus subtilis*

<400> 96

```

Ala Ile Val Lys Ala Thr Asp Gln Ser Phe Ser Ala Glu Thr Ser Glu
 1          5          10          15
Gly Val Val Leu Ala Asp Phe Trp Ala Pro Trp Cys Gly Pro Cys Lys
          20          25          30
Met Ile Ala Pro Val Leu Glu Glu Leu Asp Gln Glu Met Gly Asp Lys
          35          40          45
Leu Lys Ile Val Lys Ile Asp Val Asp Glu Asn Gln Glu Thr Ala Gly
          50          55          60
Lys Tyr Gly Val Met Ser Ile Pro Thr Leu Leu Val Leu Lys Asp Gly
65          70          75          80
Glu Val Val Glu Thr Ser Val Gly Phe Lys Pro Lys Glu Ala Leu Gln
          85          90          95
Glu Leu Val Asn Lys His Leu
          100

```

<210> 97

<211> 87

<212> PRT

<213> *Bacteriophage T4*

<400> 97

```

Met Phe Lys Val Tyr Gly Tyr Asp Ser Asn Ile His Lys Cys Val Tyr
 1          5          10          15
Cys Asp Asn Ala Lys Arg Leu Leu Thr Val Lys Lys Gln Pro Phe Glu
          20          25          30
Phe Ile Asn Ile Met Pro Glu Lys Gly Val Phe Asp Asp Glu Lys Ile
          35          40          45
Ala Glu Leu Leu Thr Lys Leu Gly Arg Asp Thr Gln Ile Gly Leu Thr
          50          55          60
Met Pro Gln Val Phe Ala Pro Asp Gly Ser His Ile Gly Gly Phe Asp
65          70          75          80
Gln Leu Arg Glu Tyr Phe Lys
          85

```

<210> 98

<211> 117

<212> PRT

<213> *Borrelia burgdorferi*

<400> 98

```

Met Ala Ile Ser Leu Thr Glu Glu Asp Phe Val Val Lys Val Phe Asp
 1          5          10          15
Tyr Lys Asn Asp Lys Glu Trp Ser Phe Arg Gly Asp Arg Pro Ala Ile
          20          25          30

```



```

Ile Asp Phe Tyr Ala Asn Trp Cys Gly Pro Cys Lys Met Leu Ser Pro
   35           40           45
Ile Phe Glu Lys Leu Ser Lys Lys Tyr Glu Asn Ser Ile Asp Phe Tyr
   50           55           60
Lys Val Asp Thr Asp Lys Glu Gln Asp Ile Ser Ser Ala Ile Gly Val
   65           70           75           80
Gln Ser Leu Pro Thr Ile Leu Phe Ile Pro Val Asp Gly Lys Pro Lys
   85           90           95
Val Ser Val Gly Phe Leu Gln Glu Asp Ala Phe Glu Asn Ile Ile Lys
   100          105          110
Asp Phe Phe Gly Phe
   115

```

<210> 99

<211> 108

<212> PRT

<213> Buchnera aphidicola

<400> 99

```

Met Asn Lys Ile Ile Glu Leu Thr Asp Gln Asn Phe Glu Glu Gln Val
  1           5           10           15
Leu Asn Ser Lys Ser Phe Phe Leu Val Asp Phe Trp Ala Gln Trp Cys
   20           25           30
Asn Pro Cys Lys Ile Leu Ala Pro Ile Leu Glu Glu Ile Ser Lys Glu
   35           40           45
Tyr Ser Asn Lys Val Ile Val Gly Lys Leu Asn Ile Glu Glu Asn Pro
   50           55           60
Asn Thr Ala Pro Val Tyr Ser Ile Arg Ser Ile Pro Thr Leu Leu Leu
   65           70           75           80
Phe Asn Asn Ser Glu Val Leu Ala Thr Lys Val Gly Ala Val Ser Lys
   85           90           95
Leu Glu Leu Lys Glu Phe Leu Asp Glu Asn Ile Asn
   100          105

```

<210> 100

<211> 108

<212> PRT

<213> aphidicola

<400> 100

```

Met Asn Lys Ile Ile Glu Leu Thr Asp Gln Asn Phe Glu Lys Glu Val
  1           5           10           15
Leu Glu His Lys Ser Phe Val Leu Val Asp Phe Trp Ala Glu Trp Cys
   20           25           30
Asn Pro Cys Lys Ile Leu Ala Pro Ile Leu Glu Glu Ile Ala Gln Glu
   35           40           45
Tyr Phe Asn Lys Ile Lys Val Gly Lys Leu Asn Ile Glu Lys Asn Pro
   50           55           60
Asn Thr Ala Pro Ile Tyr Ser Ile Arg Gly Ile Pro Ala Leu Leu Leu
   65           70           75           80
Phe His Gly Arg Glu Val Leu Ala Thr Lys Val Gly Ala Ile Ser Lys
   85           90           95
Leu Gln Leu Lys Asp Phe Leu Asp Glu Asn Ile Lys
   100          105

```

<210> 101

<211> 108

<212> PRT

<213> Chlorobium limicola

<220>

<221> VARIANT

<222> 16, 17, 38, 42, 45, 54, 55, 58, 66, 72, 75, 79, 80, 81, 94,
99, 103

<223> Xaa = Any Amino Acid

<400> 101

Ala	Gly	Lys	Tyr	Phe	Glu	Ala	Thr	Asp	Lys	Asn	Phe	Gln	Thr	Glu	Xaa
1				5					10					15	
Xaa	Asp	Ser	Asp	Lys	Ala	Val	Leu	Val	Asp	Phe	Trp	Ala	Ser	Trp	Cys
			20					25					30		
Gly	Pro	Cys	Met	Met	Xaa	Gly	Pro	Val	Xaa	Glu	Gln	Xaa	Ala	Asp	Asp
		35					40					45			
Tyr	Glu	Gly	Lys	Ala	Xaa	Xaa	Ala	Lys	Xaa	Asn	Val	Asp	Glu	Asn	Pro
	50					55					60				
Asn	Xaa	Ala	Gly	Gln	Tyr	Gly	Xaa	Arg	Ser	Xaa	Pro	Thr	Met	Xaa	Xaa
65					70					75					80
Xaa	Lys	Gly	Gly	Lys	Val	Val	Asp	Gln	Met	Val	Gly	Ala	Xaa	Pro	Lys
				85					90					95	
Asn	Met	Xaa	Ala	Lys	Lys	Xaa	Asp	Glu	His	Ile	Gly				
			100					105							

<210> 102

<211> 102

<212> PRT

<213> Chlamydia muridarum

<400> 102

Met	Val	Gln	Ile	Val	Ser	Gln	Asp	Asn	Phe	Ala	Asp	Ser	Ile	Ala	Ser
1				5					10					15	
Gly	Leu	Val	Leu	Val	Asp	Phe	Phe	Ala	Glu	Trp	Cys	Gly	Pro	Cys	Lys
			20					25					30		
Met	Leu	Thr	Pro	Val	Leu	Glu	Ala	Leu	Ala	Ala	Glu	Leu	Pro	Tyr	Val
		35					40					45			
Thr	Ile	Leu	Lys	Leu	Asp	Ile	Asp	Ala	Ser	Pro	Arg	Pro	Ala	Glu	Gln
	50					55					60				
Phe	Gly	Val	Ser	Ser	Ile	Pro	Thr	Leu	Ile	Leu	Phe	Lys	Asp	Gly	Lys
65					70					75					80
Glu	Val	Glu	Arg	Ser	Val	Gly	Leu	Lys	Asp	Lys	Asp	Ser	Leu	Val	Lys
				85					90					95	
Leu	Ile	Ser	Lys	His	Gln										
				100											

<210> 103

<211> 102

<212> PRT

<213> Chlamydia pneumoniae

<400> 103

Met	Val	Lys	Ile	Ile	Ser	Ser	Glu	Asn	Phe	Asp	Ser	Phe	Ile	Ala	Ser
1				5					10					15	
Gly	Leu	Val	Leu	Val	Asp	Phe	Phe	Ala	Glu	Trp	Cys	Gly	Pro	Cys	Arg
			20					25					30		
Met	Leu	Thr	Pro	Ile	Leu	Glu	Asn	Leu	Ala	Ala	Glu	Leu	Pro	His	Val
		35					40					45			
Thr	Ile	Gly	Lys	Ile	Asn	Ile	Asp	Glu	Asn	Ser	Lys	Pro	Ala	Glu	Thr
	50					55					60				
Tyr	Glu	Val	Ser	Ser	Ile	Pro	Thr	Leu	Ile	Leu	Phe	Lys	Asp	Gly	Asn
65					70					75					80
Glu	Val	Ala	Arg	Val	Val	Gly	Leu	Lys	Asp	Lys	Glu	Phe	Leu	Thr	Asn
				85					90					95	
Leu	Ile	Asn	Lys	His	Ala										
				100											

<210> 104

<211> 102

<212> PRT

<213> Psittaci

<400> 104
 Met Val Lys Val Val Ser Ala Glu Asn Phe Asn Ser Phe Ile Ala Thr
 1 5 10 15
 Gly Leu Val Leu Ile Asp Phe Phe Ala Glu Trp Cys Gly Pro Cys Lys
 20 25 30
 Met Leu Thr Pro Val Leu Glu Ser Leu Glu Ala Glu Val Ser Ser Val
 35 40 45
 Leu Ile Gly Lys Val Asn Ile Asp Asp His Pro Ala Pro Ala Glu Gln
 50 55 60
 Tyr Gly Val Ser Ser Ile Pro Thr Leu Ile Leu Phe Lys Asp Gly Lys
 65 70 75 80
 Glu Val Asp Arg Val Val Gly Leu Lys Asp Lys Asp Ser Leu Ile Arg
 85 90 95
 Leu Ile Asn Gln His Ser
 100

<210> 105
 <211> 102
 <212> PRT
 <213> Chlamydia trachomatis

<400> 105
 Met Val Gln Val Val Ser Gln Glu Asn Phe Ala Asp Ser Ile Ala Ser
 1 5 10 15
 Gly Leu Val Leu Ile Asp Phe Phe Ala Glu Trp Cys Gly Pro Cys Lys
 20 25 30
 Met Leu Thr Pro Val Leu Glu Ala Leu Ala Ala Glu Leu Pro His Val
 35 40 45
 Thr Ile Leu Lys Val Asp Ile Asp Ser Ser Pro Arg Pro Ala Glu Gln
 50 55 60
 Tyr Ser Val Ser Ser Ile Pro Thr Leu Ile Leu Phe Lys Asp Gly Lys
 65 70 75 80
 Glu Val Glu Arg Ser Val Gly Leu Lys Asp Lys Asp Ser Leu Ile Lys
 85 90 95
 Leu Ile Ser Lys His Gln
 100

<210> 106
 <211> 105
 <212> PRT
 <213> Corynebacterium nephridii

<400> 106
 Ala Thr Val Lys Val Asp Asn Ser Asn Phe Gln Ser Asp Val Leu Gln
 1 5 10 15
 Ser Ser Glu Pro Val Val Val Asp Phe Trp Ala Glu Trp Cys Gly Pro
 20 25 30
 Cys Lys Met Ile Ala Pro Ala Leu Asp Glu Ile Ala Thr Glu Met Ala
 35 40 45
 Gly Gln Val Lys Ile Ala Lys Val Asn Ile Asp Glu Asn Pro Glu Leu
 50 55 60
 Ala Ala Gln Phe Gly Val Arg Ser Ile Pro Thr Leu Leu Met Phe Lys
 65 70 75 80
 Asp Gly Glu Leu Ala Ala Asn Met Val Gly Ala Ala Pro Lys Ser Arg
 85 90 95
 Leu Ala Asp Trp Ile Lys Ala Ser Ala
 100 105

<210> 107
 <211> 107
 <212> PRT
 <213> Corynebacterium nephridii

<400> 107

```

Ser Ala Thr Ile Val Asn Thr Thr Asp Glu Asn Phe Gln Ala Asp Val
 1      5      10      15
Leu Asp Ala Glu Thr Pro Val Leu Val Asp Phe Trp Ala Gly Trp Cys
      20      25      30
Ala Pro Cys Lys Ala Ile Ala Pro Val Leu Glu Glu Leu Ser Asn Glu
      35      40      45
Tyr Ala Gly Lys Val Lys Ile Val Lys Val Asp Val Thr Ser Cys Glu
      50      55      60
Asp Thr Ala Val Lys Tyr Asn Ile Arg Asn Ile Pro Ala Leu Leu Met
      65      70      75      80
Phe Lys Asp Gly Glu Val Val Ala Gln Gln Val Gly Ala Ala Pro Arg
      85      90      95
Ser Lys Leu Ala Ala Phe Ile Asp Gln Asn Ile
      100      105

```

<210> 108

<211> 145

<212> PRT

<213> *Cornybacterium nephridii*

<400> 108

```

Met Ile Ile Val Cys Ala Ser Cys Gly Ala Lys Asn Arg Val Pro Glu
 1      5      10      15
Glu Lys Leu Ala Val His Pro Asn Cys Gly Gln Cys His Gln Ala Leu
      20      25      30
Leu Pro Leu Glu Pro Ile Glu Leu Asn Glu Gln Asn Phe Ser Asn Phe
      35      40      45
Ile Ser Asn Ser Asp Leu Pro Val Leu Ile Asp Leu Trp Ala Glu Trp
      50      55      60
Cys Gly Pro Cys Lys Met Met Ala Pro His Phe Ala Gln Val Ala Lys
      65      70      75      80
Gln Asn Pro Tyr Val Val Phe Ala Lys Ile Asp Thr Glu Ala Asn Pro
      85      90      95
Arg Leu Ser Ala Ala Phe Asn Val Arg Ser Ile Pro Thr Leu Val Leu
      100      105      110
Met Asn Lys Thr Thr Glu Val Ala Arg Ile Ser Gly Ala Leu Arg Thr
      115      120      125
Leu Glu Leu Gln Gln Trp Leu Asp Gln Gln Leu Gln Gln Gln Gly
      130      135      140
Asn
145

```

<210> 109

<211> 107

<212> PRT

<213> *Chromatium vinosum*

<220>

<221> VARIANT

<222> 17, 38, 42, 55, 58, 60, 72, 107

<223> Xaa = Any Amino Acid

<400> 109

```

Ser Asp Ser Ile Val His Val Thr Asp Asp Ser Phe Glu Glu Glu Val
 1      5      10      15
Xaa Lys Ser Pro Asp Pro Val Leu Val Asp Tyr Trp Ala Asp Trp Cys
      20      25      30
Gly Pro Cys Lys Met Xaa Ala Pro Val Xaa Asp Glu Ile Ala Asp Glu
      35      40      45
Tyr Ala Gly Arg Val Lys Xaa Ala Lys Xaa Asn Xaa Asp Glu Asn Pro
      50      55      60
Asn Thr Pro Pro Arg Tyr Gly Xaa Arg Gly Ile Pro Thr Leu Met Leu
      65      70      75      80
Phe Arg Gly Gly Glu Val Glu Ala Thr Lys Val Gly Ala Val Ser Lys
      85      90      95

```

Ser Gln Leu Thr Ala Phe Leu Asp Ser Asn Xaa
 100 105

<210> 110

<211> 107

<212> PRT

<213> *Clostridium litorale*

<400> 110

Met Leu Met Leu Asp Lys Asp Thr Phe Lys Thr Glu Val Leu Glu Gly
 1 5 10 15
 Thr Gly Tyr Val Leu Val Asp Tyr Phe Ser Asp Gly Cys Val Pro Cys
 20 25 30
 Lys Ala Leu Met Pro Ala Val Glu Glu Leu Ser Lys Lys Tyr Glu Gly
 35 40 45
 Arg Val Val Phe Ala Lys Leu Asn Thr Thr Gly Ala Arg Arg Leu Ala
 50 55 60
 Ile Ser Gln Lys Ile Leu Gly Leu Pro Thr Leu Ser Leu Tyr Lys Asp
 65 70 75 80
 Gly Val Lys Val Asp Glu Val Thr Lys Asp Asp Ala Thr Ile Glu Asn
 85 90 95
 Ile Glu Ala Met Val Glu Glu His Ile Ser Lys
 100 105

<210> 111

<211> 40

<212> PRT

<213> *Clostridium sporogenes*

<400> 111

Met Leu Val Leu Asp Lys Lys Thr Phe Glu Glu Glu Val Leu Lys Thr
 1 5 10 15
 Lys Gly Tyr Val Leu Val Asp Tyr Phe Gly Asp Gly Cys Val Pro Cys
 20 25 30
 Glu Ala Leu Met Pro Asp Val Glu
 35 40

<210> 112

<211> 33

<212> PRT

<213> *Clostridium sticklandii*

<400> 112

Met Phe Glu Leu Asp Lys Asp Thr Phe Glu Thr Glu Val Leu Gln Gly
 1 5 10 15
 Thr Gly Tyr Val Leu Val Asp Phe Trp Ser Glu Gly Cys Glu Pro Cys
 20 25 30
 Lys

<210> 113

<211> 106

<212> PRT

<213> *Coprinus comatus*

<400> 113

Met Val Gln Val Ile Ser Asn Leu Asp Glu Phe Asn Lys Leu Thr Asn
 1 5 10 15
 Ser Gly Lys Ile Ile Ile Ile Asp Phe Trp Ala Thr Trp Cys Gly Pro
 20 25 30
 Cys Arg Val Ile Ser Pro Ile Phe Glu Lys Phe Ser Glu Lys Tyr Gly
 35 40 45
 Ala Asn Asn Ile Val Phe Ala Lys Val Asp Val Asp Thr Ala Ser Asp

```

      50              55              60
Ile Ser Glu Glu Ala Lys Ile Arg Ala Met Pro Thr Phe Gln Val Tyr
65              70              75              80
Lys Asp Gly Gln Lys Ile Asp Glu Leu Val Gly Ala Asn Pro Thr Ala
      85              90              95
Leu Glu Ser Leu Val Gln Lys Ser Leu Ala
      100              105

```

<210> 114
 <211> 105
 <212> PRT
 <213> Dictyostelium discoideum

```

<400> 114
Met Ser Asn Arg Val Ile His Val Ser Ser Cys Glu Glu Leu Asp Lys
1              5              10              15
His Leu Arg Asp Glu Arg Val Val Val Asp Phe Ser Ala Val Trp Cys
      20              25              30
Gly Pro Cys Arg Ala Ile Ser Pro Val Phe Glu Lys Leu Ser Asn Glu
      35              40              45
Phe Ile Thr Phe Thr Phe Leu His Val Asp Ile Asp Lys Leu Asn Val
      50              55              60
His Pro Ile Val Ser Lys Ile Lys Ser Val Pro Thr Phe His Phe Tyr
65              70              75              80
Arg Asn Gly Ser Lys Val Ser Glu Phe Ser Gly Ala Ser Glu Ser Ile
      85              90              95
Leu Arg Ser Thr Leu Glu Ala Asn Lys
      100              105

```

<210> 115
 <211> 88
 <212> PRT
 <213> Dictyostelium discoideum

```

<400> 115
Met Ser Arg Val Ile His Ile Ser Ser Asn Glu Glu Leu Asp Lys His
1              5              10              15
Leu Gln Ala Glu Arg Leu Val Ile Asp Phe Ser Ala Ala Trp Cys Gly
      20              25              30
Pro Cys Arg Ala Ile Ser Pro Val Phe Glu Lys Leu Ser Asn Glu Phe
      35              40              45
Val Thr Phe Thr Phe Val His Val Asp Ile Asp Lys Leu Ser Gly His
      50              55              60
Pro Ile Val Lys Glu Ile Arg Ser Val Pro Thr Phe Tyr Phe Tyr Arg
65              70              75              80
Asn Gly Ala Lys Val Ser Glu Phe
      85

```

<210> 116
 <211> 88
 <212> PRT
 <213> Dictyostelium discoideum

```

<400> 116
Met Ser Arg Val Ile His Ile Ser Ser Asn Glu Glu Leu Asp Lys His
1              5              10              15
Leu Gln Ala Glu Arg Leu Val Ile Asp Phe Ser Ala Ala Trp Cys Gly
      20              25              30
Pro Cys Arg Ala Ile Ser Pro Val Phe Glu Lys Leu Ser Asn Glu Phe
      35              40              45
Val Thr Phe Thr Phe Val His Val Asp Ile Asp Lys Leu Ser Gly His
      50              55              60
Pro Ile Val Lys Glu Ile Arg Ser Val Pro Thr Phe Tyr Phe Tyr Arg
65              70              75              80

```

Asn Gly Ala Lys Val Ser Glu Phe
85

<210> 117

<211> 108

<212> PRT

<213> E coli, salmonella typhimurium

<400> 117

Ser	Asp	Lys	Ile	Ile	His	Leu	Thr	Asp	Asp	Ser	Phe	Asp	Thr	Asp	Val
1				5					10					15	
Leu	Lys	Ala	Asp	Gly	Ala	Ile	Leu	Val	Asp	Phe	Trp	Ala	Glu	Trp	Cys
		20						25					30		
Gly	Pro	Cys	Lys	Met	Ile	Ala	Pro	Ile	Leu	Asp	Glu	Ile	Ala	Asp	Glu
		35					40					45			
Tyr	Gln	Gly	Lys	Leu	Thr	Val	Ala	Lys	Leu	Asn	Ile	Asp	Gln	Asn	Pro
50					55					60					
Gly	Thr	Ala	Pro	Lys	Tyr	Gly	Ile	Arg	Gly	Ile	Pro	Thr	Leu	Leu	Leu
65					70				75					80	
Phe	Lys	Asn	Gly	Glu	Val	Ala	Ala	Thr	Lys	Val	Gly	Ala	Leu	Ser	Lys
			85						90					95	
Gly	Gln	Leu	Lys	Glu	Phe	Leu	Asp	Ala	Asn	Leu	Ala				
			100					105							

<210> 118

<211> 105

<212> PRT

<213> Synechocystis

<400> 118

Met	Ala	Val	Lys	Lys	Gln	Phe	Ala	Asn	Phe	Ala	Glu	Met	Leu	Ala	Gly
1				5					10					15	
Ser	Pro	Lys	Pro	Val	Leu	Val	Asp	Phe	Tyr	Ala	Thr	Trp	Cys	Gly	Pro
			20					25					30		
Cys	Gln	Met	Met	Ala	Pro	Ile	Leu	Glu	Gln	Val	Gly	Ser	His	Leu	Arg
		35					40				45				
Gln	Gln	Ile	Gln	Val	Val	Lys	Ile	Asp	Thr	Asp	Lys	Tyr	Pro	Ala	Ile
50					55					60					
Ala	Thr	Gln	Tyr	Gln	Ile	Gln	Ser	Leu	Pro	Thr	Leu	Val	Leu	Phe	Lys
65					70				75					80	
Gln	Gly	Gln	Pro	Val	His	Arg	Met	Glu	Gly	Val	Gln	Gln	Ala	Ala	Gln
			85						90					95	
Leu	Ile	Gln	Gln	Leu	Gln	Val	Phe	Val							
			100					105							

<210> 119

<211> 139

<212> PRT

<213> E. coli

<400> 119

Met	Asn	Thr	Val	Cys	Thr	His	Cys	Gln	Ala	Ile	Asn	Arg	Ile	Pro	Asp
1				5					10					15	
Asp	Arg	Ile	Glu	Asp	Ala	Ala	Lys	Cys	Gly	Arg	Cys	Gly	His	Asp	Leu
			20					25					30		
Phe	Asp	Gly	Glu	Val	Ile	Asn	Ala	Thr	Gly	Glu	Thr	Leu	Asp	Lys	Leu
		35					40				45				
Leu	Lys	Asp	Asp	Leu	Pro	Val	Val	Ile	Asp	Phe	Trp	Ala	Pro	Trp	Cys
50					55					60					
Gly	Pro	Cys	Arg	Asn	Phe	Ala	Pro	Ile	Phe	Glu	Asp	Val	Ala	Gln	Glu
65					70				75					80	
Arg	Ser	Gly	Lys	Val	Arg	Phe	Val	Lys	Val	Asn	Thr	Glu	Ala	Glu	Arg
			85						90					95	
Glu	Leu	Ser	Ser	Arg	Phe	Gly	Ile	Arg	Ser	Ile	Pro	Thr	Ile	Met	Ile

100 105 110
 Phe Lys Asn Gly Gln Val Val Asp Met Leu Asn Gly Ala Val Pro Lys
 115 120 125
 Ala Pro Phe Asp Ser Trp Leu Asn Glu Ser Leu
 130 135

<210> 120
 <211> 110
 <212> PRT
 <213> Eubacterium acidaminophilum

<400> 120
 Met Ser Ala Leu Leu Val Glu Ile Asp Lys Asp Gln Phe Gln Ala Glu
 1 5 10 15
 Val Leu Glu Ala Glu Gly Tyr Val Leu Val Asp Tyr Phe Ser Asp Gly
 20 25 30
 Cys Val Pro Cys Lys Ala Leu Met Pro Asp Val Glu Glu Leu Ala Ala
 35 40 45
 Lys Tyr Glu Gly Lys Val Ala Phe Arg Lys Phe Asn Thr Ser Ser Ala
 50 55 60
 Arg Arg Leu Ala Ile Ser Gln Lys Ile Leu Gly Leu Pro Thr Ile Thr
 65 70 75 80
 Leu Tyr Lys Gly Gly Gln Lys Val Glu Glu Val Thr Lys Asp Asp Ala
 85 90 95
 Thr Arg Glu Asn Ile Asp Ala Met Ile Ala Lys His Val Gly
 100 105 110

<210> 121
 <211> 107
 <212> PRT
 <213> Haemophilus influenzae

<400> 121
 Met Ser Glu Val Leu His Ile Asn Asp Ala Asp Phe Glu Ser Val Val
 1 5 10 15
 Val Asn Ser Asp Ile Pro Ile Leu Leu Asp Phe Trp Ala Pro Trp Cys
 20 25 30
 Gly Pro Cys Lys Met Ile Ala Pro Val Leu Asp Glu Leu Ala Pro Glu
 35 40 45
 Phe Ala Gly Lys Val Lys Ile Val Lys Met Asn Val Asp Asp Asn Gln
 50 55 60
 Ala Thr Pro Ala Gln Phe Gly Val Arg Ser Ile Pro Thr Leu Leu Leu
 65 70 75 80
 Ile Lys Asn Gly Gln Val Val Ala Thr Gln Val Gly Ala Leu Pro Lys
 85 90 95
 Thr Gln Leu Ala Asn Phe Ile Asn Gln His Ile
 100 105

<210> 122
 <211> 167
 <212> PRT
 <213> Haemophilus influenzae

<400> 122
 Met Lys Ile Lys Lys Leu Leu Lys Asn Gly Leu Ser Leu Phe Leu Thr
 1 5 10 15
 Phe Ile Val Ile Thr Ser Ile Leu Asp Phe Val Arg Arg Pro Val Val
 20 25 30
 Pro Glu Glu Ile Asn Lys Ile Thr Leu Gln Asp Leu Gln Gly Asn Thr
 35 40 45
 Phe Ser Leu Glu Ser Leu Asp Gln Asn Lys Pro Thr Leu Leu Tyr Phe
 50 55 60
 Trp Gly Thr Trp Cys Gly Tyr Cys Arg Tyr Thr Ser Pro Ala Ile Asn
 65 70 75 80

Ser Leu Ala Lys Glu Gly Tyr Gln Val Val Ser Val Ala Leu Arg Ser
 85 90 95
 Gly Asn Glu Ala Asp Val Asn Asp Tyr Leu Ser Lys Asn Asp Tyr His
 100 105 110
 Phe Thr Thr Val Asn Asp Pro Lys Gly Glu Phe Ala Glu Arg Trp Gln
 115 120 125
 Ile Asn Val Thr Pro Thr Ile Val Leu Leu Ser Lys Gly Lys Met Asp
 130 135 140
 Leu Val Thr Thr Gly Leu Thr Ser Tyr Trp Gly Leu Lys Val Arg Leu
 145 150 155 160
 Phe Phe Ala Glu Phe Phe Gly
 165

<210> 123
 <211> 106
 <212> PRT
 <213> *Helicobacter pylori*

<400> 123
 Met Ser His Tyr Ile Glu Leu Thr Glu Glu Asn Phe Glu Ser Thr Ile
 1 5 10 15
 Lys Lys Gly Val Ala Leu Val Asp Phe Trp Ala Pro Trp Cys Gly Pro
 20 25 30
 Cys Lys Met Leu Ser Pro Val Ile Asp Glu Leu Ala Ser Glu Tyr Glu
 35 40 45
 Gly Lys Ala Lys Ile Cys Lys Val Asn Thr Asp Glu Gln Glu Leu
 50 55 60
 Ser Ala Lys Phe Gly Ile Arg Ser Ile Pro Thr Leu Leu Phe Thr Lys
 65 70 75 80
 Asp Gly Glu Val Val His Gln Leu Val Gly Val Gln Thr Lys Val Ala
 85 90 95
 Leu Lys Glu Gln Leu Asn Lys Leu Leu Gly
 100 105

<210> 124
 <211> 103
 <212> PRT
 <213> *Listeria monocytogenes*

<400> 124
 Met Val Lys Glu Ile Thr Asp Ala Thr Phe Glu Gln Glu Thr Ser Glu
 1 5 10 15
 Gly Leu Val Leu Thr Asp Phe Trp Ala Thr Trp Cys Gly Pro Cys Arg
 20 25 30
 Met Val Ala Pro Val Leu Glu Glu Ile Gln Glu Glu Arg Gly Glu Ala
 35 40 45
 Leu Lys Ile Val Lys Met Asp Val Asp Glu Asn Pro Glu Thr Pro Gly
 50 55 60
 Ser Phe Gly Val Met Ser Ile Pro Thr Leu Leu Ile Lys Lys Asp Gly
 65 70 75 80
 Glu Val Val Glu Thr Ile Ile Gly Tyr Arg Pro Lys Glu Glu Leu Asp
 85 90 95
 Glu Val Ile Asn Lys Tyr Val
 100

<210> 125
 <211> 85
 <212> PRT
 <213> *Methanococcus jannaschii*

<400> 125
 Met Ser Lys Val Lys Ile Glu Leu Phe Thr Ser Pro Met Cys Pro His
 1 5 10 15
 Cys Pro Ala Ala Lys Arg Val Val Glu Glu Val Ala Asn Glu Met Pro

```

      20      25      30
Asp Ala Val Glu Val Glu Tyr Ile Asn Val Met Glu Asn Pro Gln Lys
      35      40      45
Ala Met Glu Tyr Gly Ile Met Ala Val Pro Thr Ile Val Ile Asn Gly
      50      55      60
Asp Val Glu Phe Ile Gly Ala Pro Thr Lys Glu Ala Leu Val Glu Ala
      65      70      75      80
Ile Lys Lys Arg Leu
      85

```

<210> 126
 <211> 102
 <212> PRT
 <213> Mycoplasma genitalium

```

<400> 126
Met Val Thr Glu Ile Arg Ser Leu Lys Gln Leu Glu Glu Ile Phe Ser
  1      5      10      15
Ala Lys Lys Asn Val Ile Val Asp Phe Trp Ala Ala Trp Cys Gly Pro
      20      25      30
Cys Lys Leu Thr Ser Pro Glu Phe Gln Lys Ala Ala Asp Glu Phe Ser
      35      40      45
Asp Ala Gln Phe Val Lys Val Asn Val Asp Asp His Thr Asp Ile Ala
      50      55      60
Ala Ala Tyr Asn Ile Thr Ser Leu Pro Thr Ile Val Val Phe Glu Asn
      65      70      75      80
Gly Val Glu Lys Lys Arg Ala Ile Gly Phe Met Pro Lys Thr Lys Ile
      85      90      95
Ile Asp Leu Phe Asn Asn
      100

```

<210> 127
 <211> 458
 <212> PRT
 <213> mycobacterium leprae

```

<400> 127
Met Asn Thr Thr Pro Ser Ala His Glu Thr Ile His Glu Val Ile Val
  1      5      10      15
Ile Gly Ser Gly Pro Ala Gly Tyr Thr Ala Ala Leu Tyr Ala Ala Arg
      20      25      30
Ala Gln Leu Thr Pro Leu Val Phe Glu Gly Thr Ser Phe Gly Gly Ala
      35      40      45
Leu Met Thr Thr Thr Glu Val Glu Asn Tyr Pro Gly Phe Arg Asn Gly
      50      55      60
Ile Thr Gly Pro Glu Leu Met Asp Asp Met Arg Glu Gln Ala Leu Arg
      65      70      75      80
Phe Gly Ala Glu Leu Arg Thr Glu Asp Val Glu Ser Val Ser Leu Arg
      85      90      95
Gly Pro Ile Lys Ser Val Val Thr Ala Glu Gly Gln Thr Tyr Gln Ala
      100      105      110
Arg Ala Val Ile Leu Ala Met Gly Thr Ser Val Arg Tyr Leu Gln Ile
      115      120      125
Pro Gly Glu Gln Glu Leu Leu Gly Arg Gly Val Ser Ala Cys Ala Thr
      130      135      140
Cys Asp Gly Ser Phe Phe Arg Gly Gln Asp Ile Ala Val Ile Gly Gly
      145      150      155      160
Gly Asp Ser Ala Met Glu Glu Ala Leu Phe Leu Thr Arg Phe Ala Arg
      165      170      175
Ser Val Thr Leu Val His Arg Arg Asp Glu Phe Arg Ala Ser Lys Ile
      180      185      190
Met Leu Gly Arg Ala Arg Asn Asn Asp Lys Ile Lys Phe Ile Thr Asn
      195      200      205
His Thr Val Val Ala Val Asn Gly Tyr Thr Thr Val Thr Gly Leu Arg
      210      215      220

```

Leu Arg Asn Thr Thr Thr Gly Glu Glu Thr Thr Leu Val Val Thr Gly
 225 230 235 240
 Val Phe Val Ala Ile Gly His Glu Pro Arg Ser Ser Leu Val Ser Asp
 245 250 255
 Val Val Asp Ile Asp Pro Asp Gly Tyr Val Leu Val Lys Gly Arg Thr
 260 265 270
 Thr Ser Thr Ser Met Asp Gly Val Phe Ala Ala Gly Asp Leu Val Asp
 275 280 285
 Arg Thr Tyr Arg Gln Ala Ile Thr Ala Ala Gly Ser Gly Cys Ala Ala
 290 295 300
 Ala Ile Asp Ala Glu Arg Trp Leu Ala Glu His Ala Gly Ser Lys Ala
 305 310 315 320
 Asn Glu Thr Thr Glu Glu Thr Gly Asp Val Asp Ser Thr Asp Thr Thr
 325 330 335
 Asp Trp Ser Thr Ala Met Thr Asp Ala Lys Asn Ala Gly Val Thr Ile
 340 345 350
 Glu Val Thr Asp Ala Ser Phe Phe Ala Asp Val Leu Ser Ser Asn Lys
 355 360 365
 Pro Val Leu Val Asp Phe Trp Ala Thr Trp Cys Gly Pro Cys Lys Met
 370 375 380
 Val Ala Pro Val Leu Glu Glu Ile Ala Ser Glu Gln Arg Asn Gln Leu
 385 390 395 400
 Thr Val Ala Lys Leu Asp Val Asp Thr Asn Pro Glu Met Ala Arg Glu
 405 410 415
 Phe Gln Val Val Ser Ile Pro Thr Met Ile Leu Phe Gln Gly Gly Gln
 420 425 430
 Pro Val Lys Arg Ile Val Gly Ala Lys Gly Lys Ala Ala Leu Leu Arg
 435 440 445
 Asp Leu Ser Asp Val Val Pro Asn Leu Asn
 450 455

<210> 128

<211> 102

<212> PRT

<213> Mycoplasma pneumoniae

<400> 128

Met Val Thr Glu Ile Lys Ser Leu Lys Gln Leu Gly Glu Leu Phe Ala
 1 5 10 15
 Ser Asn Asn Lys Val Ile Ile Asp Phe Trp Ala Glu Trp Cys Gly Pro
 20 25 30
 Cys Lys Ile Thr Gly Pro Glu Phe Ala Lys Ala Ala Ser Glu Val Ser
 35 40 45
 Thr Val Ala Phe Ala Lys Val Asn Val Asp Glu Gln Thr Asp Ile Ala
 50 55 60
 Ala Ala Tyr Lys Ile Thr Ser Leu Pro Thr Ile Val Leu Phe Glu Lys
 65 70 75 80
 Gly Gln Glu Lys His Arg Ala Ile Gly Phe Met Pro Lys Ala Lys Ile
 85 90 95
 Val Gln Leu Val Ser Gln
 100

<210> 129

<211> 112

<212> PRT

<213> Mycobacterium smegmatis

<400> 129

Met Ser Glu Asp Ser Ala Thr Val Ala Val Thr Asp Asp Ser Phe Ser
 1 5 10 15
 Thr Asp Val Leu Gly Ser Ser Lys Pro Val Leu Val Asp Phe Trp Ala
 20 25 30
 Thr Trp Cys Gly Pro Cys Lys Met Val Ala Pro Val Leu Glu Glu Ile
 35 40 45
 Ala Ala Glu Lys Gly Asp Gln Leu Thr Val Ala Lys Ile Asp Val Asp

50					55					60					
Val	Asp	Ala	Asn	Pro	Ala	Thr	Ala	Arg	Asp	Phe	Gln	Val	Val	Ser	Ile
65					70					75				80	
Pro	Thr	Met	Ile	Leu	Phe	Lys	Asp	Gly	Ala	Pro	Val	Lys	Arg	Ile	Val
			85						90					95	
Gly	Ala	Lys	Gly	Lys	Ala	Ala	Leu	Leu	Arg	Glu	Leu	Ser	Asp	Ala	Leu
			100					105					110		

<210> 130

<211> 115

<212> PRT

<213> Mycobacterium tuberculosis

<400> 130

Thr	Asp	Ser	Glu	Lys	Ser	Ala	Thr	Ile	Lys	Val	Thr	Asp	Ala	Ser	Phe
1				5					10					15	
Ala	Thr	Asp	Val	Leu	Ser	Ser	Asn	Lys	Pro	Val	Leu	Val	Asp	Phe	Trp
			20					25					30		
Ala	Thr	Trp	Cys	Gly	Pro	Cys	Lys	Met	Val	Ala	Pro	Val	Leu	Glu	Glu
			35				40					45			
Ile	Ala	Thr	Glu	Arg	Ala	Thr	Asp	Leu	Thr	Val	Ala	Lys	Leu	Asp	Val
	50					55				60					
Asp	Thr	Asn	Pro	Glu	Thr	Ala	Arg	Asn	Phe	Gln	Val	Val	Ser	Ile	Pro
65					70				75					80	
Thr	Leu	Ile	Leu	Phe	Lys	Asp	Gly	Gln	Pro	Val	Lys	Arg	Ile	Val	Gly
				85				90						95	
Ala	Lys	Gly	Lys	Ala	Ala	Leu	Leu	Arg	Glu	Leu	Ser	Asp	Val	Val	Pro
			100				105						110		
Asn	Leu	Asn													
															115

<210> 131

<211> 127

<212> PRT

<213> Neurospora crassa

<400> 131

Met	Ser	Asp	Gly	Val	Lys	His	Ile	Asn	Ser	Ala	Gln	Glu	Phe	Ala	Asn
1				5				10						15	
Leu	Leu	Asn	Thr	Gln	Tyr	Val	Val	Ala	Asp	Phe	Tyr	Ala	Ala	Asp	Trp
			20				25					30			
Cys	Gly	Pro	Cys	Lys	Ala	Ile	Ala	Pro	Met	Tyr	Ala	Gln	Phe	Ala	Lys
		35					40					45			
Thr	Phe	Ser	Ile	Pro	Asn	Phe	Leu	Ala	Phe	Ala	Lys	Ile	Asn	Val	Asp
	50					55				60					
Ser	Val	Gln	Gln	Val	Ala	Gln	His	Tyr	Arg	Val	Ser	Ala	Met	Pro	Thr
65					70				75					80	
Phe	Leu	Phe	Phe	Lys	Asn	Gly	Lys	Gln	Val	Ala	Val	Asn	Gly	Ser	Val
				85				90						95	
Met	Ile	Gln	Gly	Ala	Asp	Val	Asn	Ser	Leu	Arg	Ala	Ala	Ala	Glu	Lys
			100				105						110		
Met	Gly	Arg	Leu	Ala	Lys	Glu	Lys	Ala	Ala	Ala	Ala	Gly	Ser	Ser	
			115				120						125		

<210> 132

<211> 106

<212> PRT

<213> Penicillium chrysogenum

<400> 132

Met	Gly	Val	Thr	Pro	Ile	Lys	Ser	Val	Ala	Glu	Tyr	Lys	Glu	Lys	Val
1				5				10						15	
Thr	Asp	Ala	Thr	Gly	Pro	Val	Val	Val	Asp	Phe	His	Ala	Thr	Trp	Cys
			20					25						30	

Gly Pro Cys Lys Ala Ile Ala Pro Ala Leu Glu Lys Leu Ser Glu Thr
 35 40 45
 His Thr Gly Ile Gln Phe Tyr Lys Val Asp Val Asp Glu Leu Ser Glu
 50 55 60
 Val Ala Ala Ser Asn Gly Val Ser Ala Met Pro Thr Phe His Phe Tyr
 65 70 75 80
 Lys Gly Gly Glu Arg Asn Glu Glu Val Lys Gly Ala Asn Pro Ala Ala
 85 90 95
 Ile Gln Ala Gly Val Lys Ala Ile Leu Glu
 100 105

<210> 133
 <211> 108
 <212> PRT
 <213> *Pseudomonas aeruginosa*

<400> 133
 Met Ser Glu His Ile Val Asn Val Thr Asp Ala Ser Phe Glu Gln Asp
 1 5 10 15
 Val Leu Lys Ala Asp Gly Pro Val Leu Val Asp Tyr Trp Ala Glu Trp
 20 25 30
 Cys Gly Pro Cys Lys Met Ile Ala Pro Val Leu Asp Glu Val Ala Arg
 35 40 45
 Asp Tyr Gln Gly Lys Leu Lys Val Cys Lys Leu Asn Ile Asp Glu Asn
 50 55 60
 Gln Asp Thr Pro Pro Lys Tyr Gly Val Arg Gly Ile Pro Thr Leu Met
 65 70 75 80
 Leu Phe Lys Asp Gly Asn Val Glu Ala Thr Lys Val Gly Ala Leu Ser
 85 90 95
 Lys Ser Gln Leu Ala Ala Phe Leu Asp Ala Asn Ile
 100 105

<210> 134
 <211> 104
 <212> PRT
 <213> *Rhodospirillum rubrum*

<220>
 <221> VARIANT
 <222> 21, 35
 <223> Xaa = Any Amino Acid

<400> 134
 Met Lys Gln Val Ser Asp Ala Ser Phe Glu Glu Asp Val Leu Lys Ala
 1 5 10 15
 Asp Gly Pro Asn Xaa Val Asp Phe Trp Ala Glu Trp Cys Gly Pro Cys
 20 25 30
 Arg Gln Xaa Ala Pro Ala Leu Glu Glu Leu Ala Thr Ala Leu Gly Asp
 35 40 45
 Lys Val Thr Val Ala Lys Ile Asn Ile Asp Glu Asn Pro Gln Thr Pro
 50 55 60
 Ser Lys Tyr Gly Val Arg Gly Ile Pro Thr Leu Met Ile Phe Lys Asp
 65 70 75 80
 Gly Gln Val Ala Ala Thr Lys Ile Gly Ala Leu Pro Lys Thr Lys Leu
 85 90 95
 Phe Glu Trp Val Glu Ala Ser Val
 100

<210> 135
 <211> 105
 <212> PRT
 <213> *Rhodobacter sphaeroides*

<400> 135

Ser Thr Val Pro Val Thr Asp Ala Thr Phe Asp Thr Glu Val Arg Lys
 1 5 10 15
 Ser Asp Val Pro Val Val Val Asp Phe Trp Ala Glu Trp Cys Gly Pro
 20 25 30
 Cys Arg Gln Ile Gly Pro Ala Leu Glu Glu Leu Ser Lys Glu Tyr Ala
 35 40 45
 Gly Lys Val Lys Ile Val Lys Val Asn Val Asp Glu Asn Pro Glu Ser
 50 55 60
 Pro Ala Met Leu Gly Val Arg Gly Ile Pro Ala Leu Phe Leu Phe Lys
 65 70 75 80
 Asn Gly Gln Val Val Ser Asn Lys Val Gly Ala Ala Pro Lys Ala Ala
 85 90 95
 Leu Ala Thr Trp Ile Ala Ser Ala Leu
 100 105

<210> 136

<211> 130

<212> PRT

<213> Rickettsia prowazekii

<400> 136

Met Ser Cys Tyr Asn Glu Ile Thr Thr Leu Leu Glu Phe Asp Ser Asn
 1 5 10 15
 Asp Ile Asn Thr Thr Gln Arg Ile Asn Met Val Asn Asn Val Thr Asp
 20 25 30
 Ser Ser Phe Lys Asn Glu Val Leu Glu Ser Asp Leu Pro Val Met Val
 35 40 45
 Asp Phe Trp Ala Glu Trp Cys Gly Pro Cys Lys Met Leu Ile Pro Ile
 50 55 60
 Ile Asp Glu Ile Ser Lys Glu Leu Gln Asp Lys Val Lys Val Leu Lys
 65 70 75 80
 Met Asn Ile Asp Glu Asn Pro Lys Thr Pro Ser Glu Tyr Gly Ile Arg
 85 90 95
 Ser Ile Pro Thr Ile Met Leu Phe Lys Asn Gly Glu Gln Lys Asp Thr
 100 105 110
 Lys Ile Gly Leu Gln Gln Lys Asn Ser Leu Leu Asp Trp Ile Asn Lys
 115 120 125
 Ser Ile
 130

<210> 137

<211> 106

<212> PRT

<213> Streptomyces aureofaciens

<400> 137

Gly Ala Thr Val Lys Val Thr Asn Ala Thr Phe Lys Ser Asp Val Leu
 1 5 10 15
 Glu Ser Asp Lys Pro Val Leu Val His Phe Glu Gly Pro Trp Cys Gly
 20 25 30
 Pro Cys Lys Met Val Ala Pro Val Leu Asp Glu Ile Ala Asn Glu Tyr
 35 40 45
 Glu Gly Lys Val Lys Val Ala Lys Val Asn Thr Asp Glu Asn Pro Gln
 50 55 60
 Leu Ala Ser Gln Tyr Gly Val Arg Ser Ile Pro Thr Arg Leu Met Phe
 65 70 75 80
 Lys Gly Gly Glu Val Ala Ala Asn Met Val Gly Ala Ala Pro Lys Thr
 85 90 95
 Arg Leu Ala Ala Phe Leu Asp Ala Ser Leu
 100 105

<210> 138

<211> 110

<212> PRT

<213> Streptomyces coelicolor

<400> 138

```

Met Ala Gly Thr Leu Lys His Val Thr Asp Asp Ser Phe Glu Gln Asp
 1          5          10          15
Val Leu Lys Asn Asp Lys Pro Val Leu Val Asp Phe Trp Ala Ala Trp
          20          25          30
Cys Gly Pro Cys Arg Gln Ile Ala Pro Ser Leu Glu Ala Ile Ala Ala
          35          40          45
Glu Tyr Gly Asp Lys Ile Glu Ile Val Lys Leu Asn Ile Asp Glu Asn
          50          55          60
Pro Gly Thr Ala Ala Lys Tyr Gly Val Met Ser Ile Pro Thr Leu Asn
65          70          75          80
Val Tyr Gln Gly Gly Glu Val Ala Lys Thr Ile Val Gly Ala Lys Pro
          85          90          95
Lys Ala Ala Ile Val Arg Asp Leu Glu Asp Phe Ile Ala Asp
          100          105          110

```

<210> 139

<211> 107

<212> PRT

<213> Streptomyces clavuligerus

<400> 139

```

Met Ala Gly Val Leu Lys Asn Val Thr Asp Asp Thr Phe Glu Ala Asp
 1          5          10          15
Val Leu Lys Ser Glu Lys Pro Val Leu Val Asp Phe Trp Ala Glu Trp
          20          25          30
Cys Gly Pro Cys Arg Gln Ile Ala Pro Ser Leu Glu Ala Ile Thr Glu
          35          40          45
His Gly Gly Gln Ile Glu Ile Val Lys Leu Asn Ile Asp Gln Asn Pro
          50          55          60
Ala Thr Ala Ala Lys Tyr Gly Val Met Ser Ile Pro Thr Leu Asn Val
65          70          75          80
Tyr Gln Gly Gly Glu Val Val Lys Thr Ile Val Gly Ala Lys Pro Lys
          85          90          95
Ala Ala Leu Leu Arg Pro Gly Pro Val Pro Arg
          100          105

```

<210> 140

<211> 106

<212> PRT

<213> Synechocystis

<400> 140

```

Ser Ala Thr Pro Gln Val Ser Asp Ala Ser Phe Lys Glu Asp Val Leu
 1          5          10          15
Asp Ser Glu Leu Pro Val Leu Val Asp Phe Trp Ala Pro Trp Cys Gly
          20          25          30
Pro Cys Arg Met Val Ala Pro Val Asp Glu Ile Ser Gln Gln Tyr
          35          40          45
Glu Gly Lys Val Lys Val Val Lys Leu Asn Thr Asp Glu Asn Pro Asn
          50          55          60
Thr Ala Ser Gln Tyr Gly Ile Arg Ser Ile Pro Thr Leu Met Ile Phe
65          70          75          80
Lys Gly Gly Gln Arg Val Asp Met Val Val Gly Ala Val Pro Lys Thr
          85          90          95
Thr Leu Ala Ser Thr Leu Glu Lys Tyr Leu
          100          105

```

<210> 141

<211> 109

<212> PRT

<213> Synechocystis

<400> 141

```

Met Ser Leu Leu Glu Ile Thr Asp Ala Glu Phe Glu Gln Glu Thr Gln
 1          5          10          15
Gly Gln Thr Lys Pro Val Leu Val Tyr Phe Trp Ala Ser Trp Cys Gly
          20          25          30
Pro Cys Arg Leu Met Ala Pro Ala Ile Gln Ala Ile Ala Lys Asp Tyr
          35          40          45
Gly Asp Lys Leu Lys Val Leu Lys Leu Glu Val Asp Pro Asn Pro Ala
 50          55          60
Ala Val Ala Gln Cys Lys Val Glu Gly Val Pro Ala Leu Arg Leu Phe
65          70          75          80
Lys Asn Asn Glu Leu Val Met Thr His Glu Gly Ala Ile Ala Lys Pro
          85          90          95
Lys Leu Leu Glu Leu Leu Lys Glu Glu Leu Asp Phe Ile
          100          105

```

<210> 142

<211> 108

<212> PRT

<213> *Thiobacillus ferrooxidans*

<400> 142

```

Met Ser Asp Ala Ile Leu Tyr Val Ser Asp Asp Ser Phe Glu Thr Asp
 1          5          10          15
Val Leu Lys Ser Ser Lys Pro Val Leu Val Asp Phe Trp Ala Glu Trp
          20          25          30
Cys Gly Pro Cys Lys Met Ile Ala Pro Ile Leu Glu Glu Ile Ala Asp
          35          40          45
Glu Tyr Ala Asp Arg Leu Arg Val Ala Lys Phe Asn Ile Asp Glu Asn
 50          55          60
Pro Asn Thr Pro Pro Gln Tyr Ala Ile Arg Gly Ile Pro Thr Leu Leu
65          70          75          80
Leu Phe Lys Ala Gly Lys Leu Glu Ala Thr Lys Val Gly Ala Leu Ser
          85          90          95
Lys Ala Gln Leu Thr Ala Phe Leu Asp Ser Gln Leu
          100          105

```

<210> 143

<211> 91

<212> PRT

<213> *Thiocapsa roseopersicina*

<400> 143

```

Met Ser Asp Ser Ile Val His Val Thr Asp Asp Ser Phe Glu Asp Glu
 1          5          10          15
Val Leu Lys Ser Leu Glu Pro Val Leu Val Asp Tyr Trp Ala Asp Trp
          20          25          30
Cys Gly Pro Cys Lys Met Ile Ala Pro Val Leu Asp Glu Ile Ala Gly
          35          40          45
Glu Tyr Ala Gly Arg Ile Lys Val Ala Lys Leu Asn Ile Asp Glu Asn
 50          55          60
Pro Asn Thr Pro Arg Arg Tyr Gly Ile Arg Gly Ile Pro Thr Leu Met
65          70          75          80
Leu Ser Arg Gln Ser Glu Val Glu Ala Thr Lys
          85          90

```

<210> 144

<211> 44

<212> PRT

<213> *Tissierella creatinophila*

<400> 144

```

Met Ile Glu Leu Asp Lys Ser Asn Phe Glu Glu Glu Val Leu Lys Ala
 1          5          10          15

```


Glu Gly Thr Val Leu Val Asp Phe Trp Ser Pro Ser Cys Glu Pro Cys
 20 25 30
 Lys Ala Leu Met Pro His Val His Asp Phe Glu Glu
 35 40

<210> 145
 <211> 105
 <212> PRT
 <213> *Treponema pallidum*

<400> 145
 Met Ala Leu Leu Asp Ile Ser Ser Gly Asn Val Arg Lys Thr Ile Glu
 1 5 10 15
 Thr Asn Pro Leu Val Ile Val Asp Phe Trp Ala Pro Trp Cys Gly Ser
 20 25 30
 Cys Lys Met Leu Gly Pro Val Leu Glu Glu Val Glu Ser Glu Val Gly
 35 40 45
 Ser Gly Val Val Ile Gly Lys Leu Asn Val Asp Asp Asp Gln Asp Leu
 50 55 60
 Ala Val Glu Phe Asn Val Ala Ser Ile Pro Thr Leu Ile Val Phe Lys
 65 70 75 80
 Asp Gly Lys Glu Val Asp Arg Ser Ile Gly Phe Val Asp Lys Ser Lys
 85 90 95
 Ile Leu Thr Leu Ile Gln Lys Asn Ala
 100 105

<210> 146
 <211> 104
 <212> PRT
 <213> *Bos taurus*

<400> 146
 Val Lys Gln Ile Glu Ser Lys Tyr Ala Phe Gln Glu Ala Leu Asn Ser
 1 5 10 15
 Ala Gly Glu Lys Leu Val Val Val Asp Phe Ser Ala Thr Trp Cys Gly
 20 25 30
 Pro Cys Lys Met Ile Lys Pro Phe Phe His Ser Leu Ser Glu Lys Tyr
 35 40 45
 Ser Asn Val Val Phe Leu Glu Val Asp Val Asp Asp Cys Gln Asp Val
 50 55 60
 Ala Ala Glu Cys Glu Val Lys Cys Met Pro Thr Phe Gln Phe Phe Lys
 65 70 75 80
 Lys Gly Gln Lys Val Gly Glu Phe Ser Gly Ala Asn Lys Glu Lys Leu
 85 90 95
 Glu Ala Thr Ile Asn Glu Leu Ile
 100

<210> 147
 <211> 166
 <212> PRT
 <213> *Bos taurus*

<400> 147
 Met Ala Gln Arg Leu Leu Arg Arg Phe Leu Thr Ser Ile Ile Ser
 1 5 10 15
 Gly Lys Pro Ser Gln Ser Arg Trp Ala Pro Val Ala Ser Arg Ala Leu
 20 25 30
 Lys Thr Pro Gln Tyr Ser Pro Gly Tyr Leu Thr Val Thr Pro Ser Gln
 35 40 45
 Ala Arg Ser Ile Tyr Thr Thr Arg Val Cys Ser Thr Thr Phe Asn Ile
 50 55 60
 Gln Asp Gly Pro Asp Phe Gln Asp Arg Val Val Asn Ser Glu Thr Pro
 65 70 75 80
 Val Val Val Asp Phe His Ala Gln Trp Cys Gly Pro Cys Lys Ile Leu

```

      85      90      95
Gly Pro Arg Leu Glu Lys Val Val Ala Lys Gln His Gly Lys Val Val
      100      105      110
Met Ala Lys Val Asp Ile Asp Asp His Thr Asp Leu Ala Leu Glu Tyr
      115      120      125
Glu Val Ser Ala Val Pro Thr Val Leu Ala Met Lys Asn Gly Asp Val
      130      135      140
Val Asp Lys Phe Val Gly Ile Lys Asp Glu Asp Gln Leu Glu Ala Phe
145      150      155      160
Leu Lys Lys Leu Ile Gly
      165

```

<210> 148
 <211> 115
 <212> PRT
 <213> *Caenorhabditis elegans*

```

<400> 148
Met Leu Lys Arg Cys Asn Phe Lys Asn Gln Val Lys Tyr Phe Gln Ser
 1      5      10      15
Asp Phe Glu Gln Leu Ile Arg Gln His Pro Glu Lys Ile Ile Leu
      20      25      30
Asp Phe Tyr Ala Thr Trp Cys Gly Pro Cys Lys Ala Ile Ala Pro Leu
      35      40      45
Tyr Lys Glu Leu Ala Thr Thr His Lys Gly Ile Ile Phe Cys Lys Val
      50      55      60
Asp Val Asp Glu Ala Glu Asp Leu Cys Ser Lys Tyr Asp Val Lys Met
65      70      75      80
Met Pro Thr Phe Ile Phe Thr Lys Asn Gly Asp Ala Ile Glu Ala Leu
      85      90      95
Glu Gly Cys Val Glu Asp Glu Leu Arg Gln Lys Val Leu Glu His Val
      100      105      110
Ser Ala Gln
      115

```

<210> 149
 <211> 20
 <212> PRT
 <213> *Canis familiaris*

```

<400> 149
Val Lys Gln Ile Glu Phe Lys Tyr Ala Phe Gln Glu Ala Leu Asn Ser
 1      5      10      15
Ala Gly Asp Lys
      20

```

<210> 150
 <211> 104
 <212> PRT
 <213> *Gallus gallus*

```

<400> 150
Val Lys Ser Val Gly Asn Leu Ala Asp Phe Glu Ala Glu Leu Lys Ala
 1      5      10      15
Ala Gly Glu Lys Leu Val Val Val Asp Phe Ser Ala Thr Trp Cys Gly
      20      25      30
Pro Cys Lys Met Ile Lys Pro Phe Phe His Ser Leu Cys Asp Lys Phe
      35      40      45
Gly Asp Val Val Phe Ile Glu Ile Asp Val Asp Asp Ala Gln Asp Val
      50      55      60
Ala Thr His Cys Asp Val Lys Cys Met Pro Thr Phe Gln Phe Tyr Lys
65      70      75      80
Asn Gly Lys Lys Val Gln Glu Phe Ser Gly Ala Asn Lys Glu Lys Leu
      85      90      95

```

Glu Glu Thr Ile Lys Ser Leu Val
100

<210> 151

<211> 107

<212> PRT

<213> Drosophila melanogaster

<400> 151

Met	Ala	Ser	Val	Arg	Thr	Met	Asn	Asp	Tyr	His	Lys	Arg	Ile	Glu	Ala
1				5					10					15	
Ala	Asp	Asp	Lys	Leu	Ile	Val	Leu	Asp	Phe	Tyr	Ala	Thr	Trp	Cys	Gly
			20					25					30		
Pro	Cys	Lys	Glu	Met	Glu	Ser	Thr	Val	Lys	Ser	Leu	Ala	Arg	Lys	Tyr
		35					40					45			
Ser	Ser	Lys	Ala	Val	Val	Leu	Lys	Ile	Asp	Val	Asp	Lys	Phe	Glu	Glu
	50					55					60				
Leu	Thr	Glu	Arg	Tyr	Lys	Val	Arg	Ser	Met	Pro	Thr	Phe	Val	Phe	Leu
65					70					75					80
Arg	Gln	Asn	Arg	Arg	Leu	Ala	Ser	Phe	Ala	Gly	Ala	Asp	Glu	His	Lys
			85						90					95	
Leu	Thr	Asn	Met	Met	Ala	Lys	Leu	Val	Lys	Ala					
			100					105							

<210> 152

<211> 104

<212> PRT

<213> Homo sapien

<400> 152

Val	Lys	Gln	Ile	Glu	Ser	Lys	Thr	Ala	Phe	Gln	Glu	Ala	Leu	Asp	Ala
1				5					10					15	
Ala	Gly	Asp	Lys	Leu	Val	Val	Val	Asp	Phe	Ser	Ala	Thr	Trp	Cys	Gly
			20					25					30		
Pro	Cys	Lys	Met	Ile	Lys	Pro	Phe	His	Ser	Leu	Ser	Glu	Lys	Tyr	
		35					40					45			
Ser	Asn	Val	Ile	Phe	Leu	Glu	Val	Asp	Val	Asp	Asp	Cys	Gln	Asp	Val
	50					55					60				
Ala	Ser	Glu	Cys	Glu	Val	Lys	Cys	Met	Pro	Thr	Phe	Gln	Phe	Phe	Lys
65					70					75					80
Lys	Gly	Gln	Lys	Val	Gly	Glu	Phe	Ser	Gly	Ala	Asn	Lys	Glu	Lys	Leu
			85						90					95	
Glu	Ala	Thr	Ile	Asn	Glu	Leu	Val								
			100												

<210> 153

<211> 166

<212> PRT

<213> Homo sapien

<400> 153

Met	Ala	Gln	Arg	Leu	Leu	Leu	Arg	Arg	Phe	Leu	Ala	Ser	Val	Ile	Ser
1				5					10					15	
Arg	Lys	Pro	Ser	Gln	Gly	Gln	Trp	Pro	Pro	Leu	Thr	Ser	Lys	Ala	Leu
			20					25					30		
Gln	Thr	Pro	Gln	Cys	Ser	Pro	Gly	Gly	Leu	Thr	Val	Thr	Pro	Asn	Pro
		35					40					45			
Ala	Arg	Thr	Ile	Tyr	Thr	Thr	Arg	Ile	Ser	Leu	Thr	Thr	Phe	Asn	Ile
	50					55					60				
Gln	Asp	Gly	Pro	Asp	Phe	Gln	Asp	Arg	Val	Val	Asn	Ser	Glu	Thr	Pro
65					70					75					80
Val	Val	Val	Asp	Phe	His	Ala	Gln	Trp	Cys	Gly	Pro	Cys	Lys	Ile	Leu
				85					90					95	
Gly	Pro	Arg	Leu	Glu	Lys	Met	Val	Ala	Lys	Gln	His	Gly	Lys	Val	Val

```

      100      105      110
Met Ala Lys Val Asp Ile Asp Asp His Thr Asp Leu Ala Ile Glu Tyr
      115      120      125
Glu Val Ser Ala Val Pro Thr Val Leu Ala Met Lys Asn Gly Asp Val
      130      135      140
Val Asp Lys Phe Val Gly Ile Lys Asp Glu Asp Gln Leu Glu Ala Phe
145      150      155      160
Leu Lys Lys Leu Ile Gly
      165

```

<210> 154
 <211> 104
 <212> PRT
 <213> Macaca mulatta

```

<400> 154
Val Lys Gln Ile Glu Ser Lys Ala Ala Phe Gln Glu Ala Leu Asp Asp
 1      5      10      15
Ala Gly Asp Lys Leu Val Val Val Asp Phe Ser Ala Thr Trp Cys Gly
      20      25      30
Pro Cys Lys Met Ile Lys Pro Phe Phe His Ser Leu Ser Glu Lys Tyr
      35      40      45
Ser Asn Val Val Phe Leu Glu Val Asp Val Asp Asp Cys Gln Asp Val
      50      55      60
Ala Ser Glu Cys Glu Val Lys Cys Met Pro Thr Phe Gln Phe Phe Lys
65      70      75      80
Lys Gly Gln Lys Val Gly Glu Phe Ser Gly Ala Asn Lys Glu Lys Leu
      85      90      95
Glu Ala Thr Ile Asn Glu Leu Val
      100

```

<210> 155
 <211> 104
 <212> PRT
 <213> Mus musculus

```

<400> 155
Val Lys Leu Ile Glu Ser Lys Glu Ala Phe Gln Glu Ala Leu Ala Ala
 1      5      10      15
Ala Gly Asp Lys Leu Val Val Val Asp Phe Ser Ala Thr Trp Cys Gly
      20      25      30
Pro Cys Lys Met Ile Lys Pro Phe Phe His Ser Leu Cys Asp Lys Tyr
      35      40      45
Ser Asn Val Val Phe Leu Glu Val Asp Val Asp Asp Cys Gln Asp Val
      50      55      60
Ala Ala Asp Cys Glu Val Lys Cys Met Pro Thr Phe Gln Phe Tyr Lys
65      70      75      80
Lys Gly Gln Lys Val Gly Glu Phe Ser Gly Ala Asn Lys Glu Lys Leu
      85      90      95
Glu Ala Ser Ile Thr Glu Tyr Ala
      100

```

<210> 156
 <211> 166
 <212> PRT
 <213> Mus musculus

```

<400> 156
Met Ala Gln Arg Leu Leu Leu Gly Arg Phe Leu Thr Ser Val Ile Ser
 1      5      10      15
Arg Lys Pro Pro Gln Gly Val Trp Ala Ser Leu Thr Ser Lys Thr Leu
      20      25      30
Gln Thr Pro Gln Tyr Asn Ala Gly Gly Leu Thr Val Met Pro Ser Pro
      35      40      45

```

Ala Arg Thr Val His Thr Thr Arg Val Cys Leu Thr Thr Phe Asn Val
 50 55 60
 Gln Asp Gly Pro Asp Phe Gln Asp Arg Val Val Asn Ser Glu Thr Pro
 65 70 75 80
 Val Val Val Asp Phe His Ala Gln Trp Cys Gly Pro Cys Lys Ile Leu
 85 90 95
 Gly Pro Arg Leu Glu Lys Met Val Ala Lys Gln His Gly Lys Val Val
 100 105 110
 Met Ala Lys Val Asp Ile Asp Asp His Thr Asp Leu Ala Ile Glu Tyr
 115 120 125
 Glu Val Ser Ala Val Pro Thr Val Leu Ala Ile Lys Asn Gly Asp Val
 130 135 140
 Val Asp Lys Phe Val Gly Ile Lys Asp Glu Asp Gln Leu Glu Ala Phe
 145 150 155 160
 Leu Lys Lys Leu Ile Gly
 165

<210> 157
 <211> 33
 <212> PRT
 <213> Sus scrofa

<400> 157
 Val Lys Gln Ile Glu Ser Lys Tyr Ala Phe Gln Glu Ala Leu Asn Ser
 1 5 10 15
 Ala Gly Glu Lys Leu Val Val Val Asp Phe Ser Ala Thr Trp Cys Gly
 20 25 30
 Pro

<210> 158
 <211> 104
 <212> PRT
 <213> Oryctolagus cuniculus

<400> 158
 Val Lys Gln Ile Glu Ser Lys Ser Ala Phe Gln Glu Val Leu Asp Ser
 1 5 10 15
 Ala Gly Asp Lys Leu Val Val Val Asp Phe Ser Ala Thr Trp Cys Gly
 20 25 30
 Pro Cys Lys Met Ile Lys Pro Phe His Ala Leu Ser Glu Lys Phe
 35 40 45
 Asn Asn Val Val Phe Ile Glu Val Asp Val Asp Asp Cys Lys Asp Ile
 50 55 60
 Ala Ala Glu Cys Glu Val Lys Cys Met Pro Thr Phe Gln Phe Phe Lys
 65 70 75 80
 Lys Gly Gln Lys Val Gly Glu Phe Ser Gly Ala Asn Lys Glu Lys Leu
 85 90 95
 Glu Ala Thr Ile Asn Glu Leu Leu
 100

<210> 159
 <211> 104
 <212> PRT
 <213> Rattus norvegicus

<400> 159
 Val Lys Leu Ile Glu Ser Lys Glu Ala Phe Gln Glu Ala Leu Ala Ala
 1 5 10 15
 Ala Gly Asp Lys Leu Val Val Val Asp Phe Ser Ala Thr Trp Cys Gly
 20 25 30
 Pro Cys Lys Met Ile Lys Pro Phe Phe His Ser Leu Cys Asp Lys Tyr
 35 40 45
 Ser Asn Val Val Phe Leu Glu Val Asp Val Asp Asp Cys Gln Asp Val

50 55 60
 Ala Ala Asp Cys Glu Val Lys Cys Met Pro Thr Phe Gln Phe Tyr Lys
 65 70 75 80
 Lys Gly Gln Lys Val Gly Glu Phe Ser Gly Ala Asn Lys Glu Lys Leu
 85 90 95
 Glu Ala Thr Ile Thr Glu Phe Ala
 100

<210> 160
 <211> 166
 <212> PRT
 <213> *Rattus norvegicus*

<400> 160
 Met Ala Gln Arg Leu Leu Arg Arg Phe Leu Thr Ser Val Ile Ser
 1 5 10 15
 Arg Lys Pro Pro Gln Gly Val Trp Ala Ser Leu Thr Ser Thr Ser Leu
 20 25 30
 Gln Thr Pro Pro Tyr Asn Ala Gly Gly Leu Thr Gly Thr Pro Ser Pro
 35 40 45
 Ala Arg Thr Phe His Thr Thr Arg Val Cys Ser Thr Thr Phe Asn Val
 50 55 60
 Gln Asp Gly Pro Asp Phe Gln Asp Arg Val Val Asn Ser Glu Thr Pro
 65 70 75 80
 Val Val Val Asp Phe His Ala Gln Trp Cys Gly Pro Cys Lys Ile Leu
 85 90 95
 Gly Pro Arg Leu Glu Lys Met Val Ala Lys Gln His Gly Lys Val Val
 100 105 110
 Met Ala Lys Val Asp Ile Asp Asp His Thr Asp Leu Ala Ile Glu Tyr
 115 120 125
 Glu Val Ser Ala Val Pro Thr Val Leu Ala Ile Lys Asn Gly Asp Val
 130 135 140
 Val Asp Lys Phe Val Gly Ile Lys Asp Glu Asp Gln Leu Glu Ala Phe
 145 150 155 160
 Leu Lys Lys Leu Ile Gly
 165

<210> 161
 <211> 104
 <212> PRT
 <213> *Ovis aries*

<400> 161
 Val Lys Gln Ile Glu Ser Lys Tyr Ala Phe Gln Glu Ala Leu Asn Ser
 1 5 10 15
 Ala Gly Glu Lys Leu Val Val Val Asp Phe Ser Ala Thr Trp Cys Gly
 20 25 30
 Pro Cys Lys Met Ile Lys Pro Phe Phe His Ser Leu Ser Glu Lys Tyr
 35 40 45
 Ser Asn Val Val Phe Leu Glu Val Asp Val Asp Asp Cys Gln Asp Val
 50 55 60
 Ala Ala Glu Cys Glu Val Lys Cys Met Pro Thr Phe Gln Phe Phe Lys
 65 70 75 80
 Lys Gly Gln Lys Val Ser Glu Phe Ser Gly Ala Asn Lys Glu Lys Leu
 85 90 95
 Glu Ala Thr Ile Asn Glu Leu Ile
 100

<210> 162
 <211> 261
 <212> PRT
 <213> *Arabidopsis thaliana*

<400> 162

Met Ala Arg Leu Val Phe Ser Leu Asn Leu Pro Ser Ser His Gly Phe
 1 5 10 15
 Asn Leu Ser Pro Arg Asn Leu Gln Ser Phe Phe Val Thr Gln Thr Gly
 20 25 30
 Ala Pro Arg Phe Arg Ala Val Arg Cys Lys Pro Asn Pro Glu Ser Ser
 35 40 45
 Glu Thr Lys Gln Glu Lys Leu Val Ile Asp Asn Gly Glu Thr Ser Ser
 50 55 60
 Ala Ser Lys Glu Val Glu Ser Ser Ser Ser Val Ala Asp Ser Ser Ser
 65 70 75 80
 Ser Ser Ser Ser Gly Phe Pro Glu Ser Pro Asn Lys Asp Ile Asn Arg
 85 90 95
 Arg Val Ala Ala Val Thr Val Ile Ala Ala Leu Ser Leu Phe Val Ser
 100 105 110
 Thr Arg Leu Asp Phe Gly Ile Ser Leu Lys Asp Leu Thr Ala Ser Ala
 115 120 125
 Leu Pro Tyr Glu Glu Ala Leu Ser Asn Gly Lys Pro Thr Val Val Glu
 130 135 140
 Phe Tyr Ala Asp Trp Cys Glu Val Cys Arg Glu Leu Ala Pro Asp Val
 145 150 155 160
 Tyr Lys Ile Glu Gln Gln Tyr Lys Asp Lys Val Asn Phe Val Met Leu
 165 170 175
 Asn Val Asp Asn Thr Lys Trp Glu Gln Glu Leu Asp Glu Phe Gly Val
 180 185 190
 Glu Gly Ile Pro His Phe Ala Phe Leu Asp Arg Glu Gly Asn Glu Glu
 195 200 205
 Gly Asn Val Val Gly Arg Leu Pro Arg Gln Tyr Leu Val Glu Asn Val
 210 215 220
 Asn Ala Leu Ala Ala Gly Lys Gln Ser Ile Pro Tyr Ala Arg Ala Val
 225 230 235 240
 Gly Gln Tyr Ser Ser Ser Glu Ser Arg Lys Val His Gln Val Thr Asp
 245 250 255
 Pro Leu Ser His Gly
 260

<210> 163

<211> 140

<212> PRT

<213> Arabidopsis thaliana

<400> 163

Met Gly Ser Cys Val Ser Lys Gly Lys Gly Asp Asp Asp Ser Val His
 1 5 10 15
 Asn Val Glu Phe Ser Gly Gly Asn Val His Leu Ile Thr Thr Lys Glu
 20 25 30
 Ser Trp Asp Asp Lys Leu Ala Glu Ala Asp Arg Asp Gly Lys Ile Val
 35 40 45
 Val Ala Asn Phe Ser Ala Thr Trp Cys Gly Pro Cys Lys Ile Val Ala
 50 55 60
 Pro Phe Phe Ile Glu Leu Ser Glu Lys His Ser Ser Leu Met Phe Leu
 65 70 75 80
 Leu Val Asp Val Asp Glu Leu Ser Asp Phe Ser Ser Ser Trp Asp Ile
 85 90 95
 Lys Ala Thr Pro Thr Phe Phe Phe Leu Lys Asn Gly Gln Gln Ile Gly
 100 105 110
 Lys Leu Val Gly Ala Asn Lys Pro Glu Leu Gln Lys Lys Val Thr Ser
 115 120 125
 Ile Ile Asp Ser Val Pro Glu Ser Pro Gln Arg Pro
 130 135 140

<210> 164

<211> 186

<212> PRT

<213> Arabidopsis thaliana

<400> 164

```

Met Ser Glu Ile Val Asn Leu Ser Ser Ser Leu Arg Ser Leu Asn Pro
 1          5          10          15
Lys Ile Ser Pro Leu Val Pro Pro Tyr Arg Gln Thr Ser Ser Ser Phe
          20          25          30
Ser Arg Pro Arg Asn Phe Lys Tyr His Ser Phe Thr Asp Lys Ile Cys
          35          40          45
Leu Ala Ala Glu Arg Ile Arg Ala Val Asp Ile Gln Lys Gln Asp Gly
          50          55          60
Gly Leu Gln Glu Leu Asp Asp Ser Pro Val Ser Val Glu Leu Gly Pro
65          70          75          80
Ile Cys Gly Glu Ser His Phe Asp Gln Val Met Glu Asp Ala Gln Lys
          85          90          95
Leu Gly Glu Ser Val Val Ile Val Trp Met Ala Ala Trp Cys Arg Lys
          100          105          110
Cys Ile Tyr Leu Lys Pro Lys Leu Glu Lys Leu Ala Ala Glu Phe Tyr
          115          120          125
Pro Arg Leu Arg Phe Tyr His Val Asp Val Asn Ala Val Pro Tyr Arg
          130          135          140
Leu Val Ser Arg Ala Gly Val Thr Leu Trp Arg Asp Gly Gln Lys Gln
145          150          155          160
Ala Glu Val Ile Gly Gly His Lys Ala His Phe Val Val Asn Glu Val
          165          170          175
Arg Glu Met Ile Glu Asn Asp Ser Ile Thr
          180          185

```

<210> 165

<211> 207

<212> PRT

<213> Arabidopsis thaliana

<400> 165

```

Met Glu Asn Met Ser Asn Leu Thr Ser Lys Phe Leu Leu Asn Pro Leu
 1          5          10          15
Asn Val His Lys His Cys Ala Val Ser Asp Glu Asn Gly Asp Arg Lys
          20          25          30
Ser His Val Leu Lys Gln Val Cys Ser Cys Ile Cys Cys Cys Asn Arg
          35          40          45
Arg Asn Lys Thr Gln Ala Arg Ser Gln Lys Gly Ser Tyr Phe Ile Lys
          50          55          60
Gly Lys Val His Pro Val Ser Arg Met Glu Lys Trp Glu Glu Lys Ile
65          70          75          80
Thr Glu Ala Asn Ser His Gly Lys Ile Ile Ala Arg His Asp Leu Ile
          85          90          95
Leu Cys Asn Met Glu Gln Leu Val Val Asn Phe Lys Ala Ser Trp Cys
          100          105          110
Leu Pro Ser Lys Thr Ile Leu Pro Ile Tyr Gln Glu Leu Ala Ser Thr
          115          120          125
Tyr Thr Ser Met Ile Phe Val Thr Ile Asp Val Glu Glu Leu Ala Ile
          130          135          140
Ser Lys Leu Ser Asp Leu Gly Val Lys Ile Cys Leu Ile Gln Glu Phe
145          150          155          160
Ser His Glu Trp Asn Val Asp Ala Thr Pro Thr Val Val Phe Leu Lys
          165          170          175
Asp Gly Arg Gln Met Asp Lys Leu Val Gly Gly Asp Ala Ala Glu Leu
          180          185          190
Gln Lys Lys Thr Ala Ala Ala Ala Asn Leu Leu Leu Arg Gln Ser
          195          200          205

```

<210> 166

<211> 175

<212> PRT

<213> Arabidopsis thaliana

<400> 166


```

Met Leu Ile Pro His Ala Val Ser Phe Ala Phe Thr Tyr Leu Arg Asn
1      5      10      15
Ser Ala Asn Pro Asp Gln Asn Arg Glu Val Ile Ser Ile His Ser Thr
20      25      30
Ser Glu Leu Glu Ala Lys Thr Lys Ala Ala Lys Lys Ala Ser Arg Leu
35      40      45
Leu Ile Leu Tyr Phe Thr Ala Thr Trp Cys Gly Pro Cys Arg Tyr Met
50      55      60
Ser Pro Leu Tyr Ser Asn Leu Ala Thr Gln His Ser Arg Val Val Phe
65      70      75      80
Leu Lys Val Asp Ile Asp Lys Ala Asn Asp Val Ala Ala Ser Trp Asn
85      90      95
Ile Ser Ser Val Pro Thr Phe Cys Phe Ile Arg Asp Gly Lys Glu Val
100      105      110
Asp Lys Val Val Gly Ala Asp Lys Gly Ser Leu Glu Gln Lys Ile Ala
115      120      125
Gln His Ser Ser Ser Lys Ala Arg Tyr Ile Pro Val Phe Ile Lys Tyr
130      135      140
His Ser Asp Leu Leu Leu Leu Val Asn Glu Glu Thr Pro Thr Ser Asn
145      150      155      160
Gln Lys Leu Lys Thr Lys Thr Gly Asp Trp Phe His Ile Asn Leu
165      170      175

```

<210> 167

<211> 132

<212> PRT

<213> Arabidopsis thaliana

<400> 167

```

Met Arg Lys Gln Glu Ser Glu Gly Ala Asn Leu Glu Phe Glu Ser Lys
1      5      10      15
Ser Asn Asp Asn Gly Asn Val Lys Ile Ala Pro Asn Asp Gln Ser Phe
20      25      30
Leu Thr Ile Leu Asp Asp Ile Lys Ser Ser Lys Ser Pro Ala Val Ile
35      40      45
Asn Tyr Gly Ala Ser Trp Tyr Thr Leu Phe Ser Val Phe Thr Ile Thr
50      55      60
Leu Phe Met Leu Ile Lys Cys Ser Met Lys Cys Leu Asn Glu Asn Gly
65      70      75      80
Phe Val Leu Lys Leu Ser Asp Ile Asp Glu Cys Pro Glu Thr Thr Arg
85      90      95
His Ile Arg Tyr Thr Pro Thr Phe Gln Phe Tyr Arg Asp Gly Glu Lys
100      105      110
Val Asp Glu Met Phe Gly Ala Gly Glu Gln Arg Leu His Asp Arg Leu
115      120      125
Trp Leu His Ser
130

```

<210> 168

<211> 151

<212> PRT

<213> Arabidopsis thaliana

<400> 168

```

Met Ala Ser Ile Ser Leu Ser Ser Ser Thr Val Pro Ser Leu Asn Ser
1      5      10      15
Lys Glu Ser Ser Gly Val Ser Ala Phe Ala Ser Arg Ser Ile Ser Ala
20      25      30
Val Lys Phe Gln Phe Pro Val Arg Arg Ile Glu Ala Lys Lys Gln Thr
35      40      45
Phe Asp Ser Phe Glu Asp Leu Leu Val Asn Ser Asp Lys Pro Val Leu
50      55      60
Val Asp Tyr Tyr Ala Thr Trp Cys Gly Pro Cys Gln Phe Met Val Pro
65      70      75      80
Ile Leu Asn Glu Val Ser Glu Thr Leu Lys Asp Lys Ile Gln Val Val

```

85 90 95
 Lys Ile Asp Thr Glu Lys Tyr Pro Ser Ile Ala Asn Lys Tyr Lys Ile
 100 105 110
 Glu Ala Leu Pro Thr Phe Ile Leu Phe Lys Asp Gly Glu Pro Cys Asp
 115 120 125
 Arg Phe Glu Gly Ala Leu Thr Ala Lys Gln Leu Ile Gln Arg Ile Glu
 130 135 140
 Asp Ser Leu Lys Val Lys Pro
 145 150

<210> 169

<211> 236

<212> PRT

<213> Arabidopsis thaliana

<400> 169

Met Ala Gly Val Val Arg Leu Thr Thr Thr Ser Val Gln Ala Ile Arg
 1 5 10 15
 Val Ser Ser Ser Phe Ser Ser Phe Ala Thr Ala Leu Asn Pro Leu Gln
 20 25 30
 Pro Cys Leu Pro Pro Asn Ser Asn Leu Asn Ser Asp Lys Arg Leu Arg
 35 40 45
 Leu Leu Ser Ser Ser Pro Ser Cys Ser Ser Ser His Tyr His Pro Ser
 50 55 60
 Ser Gly Leu Gly Ser His Leu Pro Leu Arg Arg Pro Lys Ser Gln Val
 65 70 75 80
 Val Arg Val Lys Val Asp Glu Asn Val Ala Glu Thr Glu Pro Pro Lys
 85 90 95
 Trp Trp Glu Arg Asn Ala Pro Asn Met Val Asp Ile His Ser Thr Glu
 100 105 110
 Glu Phe Leu Ser Ala Leu Ser Gly Ala Gly Glu Arg Leu Val Ile Val
 115 120 125
 Glu Phe Tyr Gly Thr Trp Cys Ala Ser Cys Arg Ala Leu Phe Pro Lys
 130 135 140
 Leu Cys Lys Thr Ala Val Glu His Pro Asp Ile Val Phe Leu Lys Val
 145 150 155 160
 Asn Phe Asp Glu Asn Lys Pro Met Cys Lys Ser Leu Asn Val Arg Val
 165 170 175
 Leu Pro Phe Phe His Phe Tyr Arg Gly Ala Asp Gly Gln Leu Glu Ser
 180 185 190
 Phe Ser Cys Ser Leu Ala Lys Val Lys Lys Ala Ile Ser Val Ser Pro
 195 200 205
 Phe Pro Gln Leu Glu Leu Gly Ile Thr Leu Gln Thr Lys Arg Thr Thr
 210 215 220
 Ser Leu Phe Phe Phe Asp Arg Ile Tyr Gln Ile Leu
 225 230 235

<210> 170

<211> 131

<212> PRT

<213> Hordeum bulbosum

<400> 170

Met Gly Gly Cys Val Gly Lys Asp Arg Ser Ile Val Glu Asp Lys Leu
 1 5 10 15
 Asp Phe Lys Gly Asn Val His Val Ile Thr Thr Lys Glu Asp Trp
 20 25 30
 Asp Gln Lys Val Ala Glu Ala Asn Lys Asp Gly Lys Ile Val Val Ala
 35 40 45
 Asn Phe Ser Ala Ser Trp Cys Gly Pro Cys Arg Val Ile Ala Pro Val
 50 55 60
 Tyr Ala Glu Met Ser Lys Thr Tyr Pro Gln Leu Met Phe Leu Thr Ile
 65 70 75 80
 Asp Val Asp Asp Leu Met Asp Phe Gly Ser Thr Trp Asp Ile Arg Ala
 85 90 95

Thr Pro Thr Phe Phe Phe Leu Lys Asn Gly Gln Gln Ile Asp Lys Leu
 100 105 110
 Val Gly Ala Asn Lys Pro Glu Leu Glu Lys Lys Val Gln Ala Leu Gly
 115 120 125
 Asp Gly Ser
 130

<210> 171
 <211> 131
 <212> PRT
 <213> *Lolium perenne*

<400> 171
 Met Gly Gly Cys Val Gly Lys Asp Arg Ser Ile Val Glu Asp Lys Leu
 1 5 10 15
 Asp Phe Lys Gly Gly Asn Val His Val Ile Thr Thr Lys Glu Asp Trp
 20 25 30
 Asp Gln Lys Val Ala Glu Ala Asn Lys Asp Gly Lys Ile Val Val Ala
 35 40 45
 Asn Phe Ser Ala Ser Trp Cys Gly Pro Cys Arg Val Ile Ala Pro Val
 50 55 60
 Tyr Ala Glu Met Ser Lys Thr Tyr Pro Gln Leu Met Phe Leu Thr Ile
 65 70 75 80
 Asp Val Asp Asp Leu Met Asp Phe Ser Ser Thr Trp Asp Ile Arg Ala
 85 90 95
 Thr Pro Thr Phe Phe Leu Lys Asn Gly Gln Leu Ile Asp Lys Leu
 100 105 110
 Val Gly Ala Asn Arg Pro Glu Leu Glu Lys Lys Val Gln Ala Ile Gly
 115 120 125
 Asp Gly Ser
 130

<210> 172
 <211> 131
 <212> PRT
 <213> *Oryza sativa*

<400> 172
 Met Gly Ser Cys Val Gly Lys Glu Arg Ser Asp Glu Glu Asp Lys Ile
 1 5 10 15
 Asp Phe Lys Gly Gly Asn Val His Val Ile Ser Asn Lys Glu Asn Trp
 20 25 30
 Asp His Lys Ile Ala Glu Ala Asn Lys Asp Gly Lys Ile Val Ile Ala
 35 40 45
 Asn Phe Ser Ala Ala Trp Cys Gly Pro Cys Arg Val Ile Ala Pro Val
 50 55 60
 Tyr Ala Glu Met Ser Gln Thr Tyr Pro Gln Phe Met Phe Leu Thr Ile
 65 70 75 80
 Asp Val Asp Glu Leu Met Asp Phe Ser Ser Ser Trp Asp Ile Arg Ala
 85 90 95
 Thr Pro Thr Phe Phe Leu Lys Asn Gly Glu Gln Val Asp Lys Leu
 100 105 110
 Val Gly Ala Asn Lys Pro Glu Leu Glu Lys Lys Val Ala Ala Leu Ala
 115 120 125
 Asp Ser Ala
 130

<210> 173
 <211> 296
 <212> PRT
 <213> *Solanum tuberosum*

<400> 173
 Met Ala Thr Leu Thr Asn Phe Leu Leu Lys Pro Ser Pro Asn Leu Ala

```

1           5           10           15
Ser Ile Thr Lys Ile Ser Pro Ser Leu Tyr Ser Asn Phe Pro Phe Glu
20           25           30
Lys Ser Lys Gln Ser Ile Phe Lys Asn Leu Lys Thr Asn Lys Pro Leu
35           40           45
Leu Ile Thr Lys Ala Thr Ala Pro Asp Val Glu Lys Lys Val Ala
50           55           60
Lys Ser Glu Arg Val Gln Lys Val Asn Ser Met Glu Glu Leu Asp Glu
65           70           75           80
Ala Leu Lys Lys Ala Lys Asn Arg Leu Val Val Val Glu Phe Ala Gly
85           90           95
Lys Asp Ser Glu Arg Ser Lys Asn Ile Tyr Pro Phe Met Val Asn Leu
100          105          110
Ser Lys Thr Cys Asn Asp Val Asp Phe Leu Leu Val Ile Gly Asp Glu
115          120          125
Thr Glu Lys Thr Lys Ala Leu Cys Arg Arg Glu Lys Ile Asp Lys Val
130          135          140
Pro His Phe Asn Phe Tyr Lys Ser Met Glu Lys Ile His Glu Glu Glu
145          150          155          160
Gly Ile Gly Pro Asp Leu Leu Ala Gly Asp Val Leu Tyr Tyr Gly Asp
165          170          175
Ser His Ser Glu Val Val Gln Leu His Ser Arg Glu Asp Val Glu Lys
180          185          190
Val Ile Gln Asp His Lys Ile Asp Lys Lys Leu Ile Val Leu Asp Val
195          200          205
Gly Leu Lys His Cys Gly Pro Cys Val Lys Val Tyr Pro Thr Val Ile
210          215          220
Lys Leu Ser Lys Gln Met Ala Asp Thr Val Val Phe Ala Arg Met Asn
225          230          235          240
Gly Asp Glu Asn Asp Ser Cys Met Gln Phe Leu Lys Asp Met Asp Val
245          250          255
Ile Glu Val Pro Thr Phe Leu Phe Ile Arg Asp Gly Glu Ile Cys Gly
260          265          270
Arg Tyr Val Gly Ser Gly Lys Gly Glu Leu Ile Gly Glu Ile Leu Arg
275          280          285
Tyr Gln Gly Val Arg Val Thr Tyr
290          295

```

<210> 174

<211> 131

<212> PRT

<213> Secale cereale

<400> 174

```

Met Gly Gly Cys Val Gly Lys Gly Arg Ser Ile Val Glu Glu Lys Leu
1           5           10           15
Asp Phe Lys Gly Gly Asn Val His Val Ile Thr Thr Lys Glu Asp Trp
20           25           30
Asp Gln Lys Ile Glu Glu Ala Asn Lys Asp Gly Lys Ile Val Val Ala
35           40           45
Asn Phe Ser Ala Ser Trp Cys Gly Pro Cys Arg Val Val Ala Pro Val
50           55           60
Tyr Ala Gly Met Ser Lys Thr Tyr Pro Gln Leu Met Phe Leu Thr Ile
65           70           75           80
Asp Val Asp Asp Leu Met Asp Phe Ser Ser Thr Trp Asp Ile Arg Ala
85           90           95
Thr Pro Thr Phe Phe Phe Leu Lys Asn Gly Gln Gln Ile Asp Lys Leu
100          105          110
Val Gly Ala Asn Lys Pro Glu Leu Glu Lys Lys Val Gln Ala Leu Gly
115          120          125
Asp Gly Ser
130

```

<210> 175

<211> 119

<212> PRT

<213> Secale cereale

<400> 175

```

Met Gly Gly Cys Val Gly Lys Gly Arg Ser Ile Val Glu Glu Lys Leu
 1          5          10
Asp Phe Lys Gly Gly Asn Val His Val Ile Thr Thr Lys Glu Asp Trp
          20          25          30
Asp Gln Lys Ile Glu Glu Ala Asn Lys Asp Gly Lys Ile Val Val Ala
          35          40          45
Asn Phe Ser Ala Ser Trp Cys Gly Pro Cys Arg Val Ile Ala Pro Val
          50          55          60
Tyr Ala Glu Met Ser Lys Thr Tyr Pro Gln Leu Met Phe Leu Thr Ile
65          70          75          80
Asp Val Asp Asp Leu Met Asp Phe Ser Ser Thr Trp Asp Ile Arg Ala
          85          90          95
Thr Pro Thr Phe Phe Leu Lys Asn Gly Gln Gln Ile Asp Lys Leu
          100          105          110
Val Gly Ala Asn Lys Pro Glu
          115

```

<210> 176

<211> 106

<212> PRT

<213> Manduca sexta

<400> 176

```

Met Ser Ile His Ile Lys Asp Ala Asp Asp Leu Lys Asn Arg Leu Ala
 1          5          10          15
Glu Ala Gly Asp Lys Leu Val Val Ile Asp Phe Met Ala Thr Trp Cys
          20          25          30
Gly Pro Cys Lys Met Ile Gly Pro Lys Leu Asp Glu Met Ala Ala Glu
          35          40          45
Met Ala Asp Ser Ile Val Val Val Lys Val Asp Val Asp Glu Cys Glu
          50          55          60
Asp Ile Ala Ala Asp Tyr Asn Ile Asn Ser Met Pro Thr Phe Val Phe
65          70          75          80
Val Lys Asn Ser Lys Lys Leu Glu Glu Phe Ser Gly Ala Asn Val Asp
          85          90          95
Lys Leu Lys Asn Thr Ile Leu Lys Leu Lys
          100          105

```

<210> 177

<211> 221

<212> PRT

<213> Bradyrhizobium japonicum

<400> 177

```

Met Leu Asp Thr Lys Pro Ser Ala Thr Arg Arg Ile Pro Leu Val Ile
 1          5          10          15
Ala Thr Val Ala Val Gly Gly Leu Ala Gly Phe Ala Ala Leu Tyr Gly
          20          25          30
Leu Gly Leu Ser Arg Ala Pro Thr Gly Asp Pro Ala Cys Arg Ala Ala
          35          40          45
Val Ala Thr Ala Gln Lys Ile Ala Pro Leu Ala His Gly Glu Val Ala
          50          55          60
Ala Leu Thr Met Ala Ser Ala Pro Leu Lys Leu Pro Asp Leu Ala Phe
65          70          75          80
Glu Asp Ala Asp Gly Lys Pro Lys Lys Leu Ser Asp Phe Arg Gly Lys
          85          90          95
Thr Leu Leu Val Asn Leu Trp Ala Thr Trp Cys Val Pro Cys Arg Lys
          100          105          110
Glu Met Pro Ala Leu Asp Glu Leu Gln Gly Lys Leu Ser Gly Pro Asn
          115          120          125
Phe Glu Val Val Ala Ile Asn Ile Asp Thr Arg Asp Pro Glu Lys Pro

```

```

      130              135              140
Lys Thr Phe Leu Lys Glu Ala Asn Leu Thr Arg Leu Gly Tyr Phe Asn
145              150              155              160
Asp Gln Lys Ala Lys Val Phe Gln Asp Leu Lys Ala Ile Gly Arg Ala
      165              170              175
Leu Gly Met Pro Thr Ser Val Leu Val Asp Pro Gln Gly Cys Glu Ile
      180              185              190
Ala Thr Ile Ala Gly Pro Ala Glu Trp Ala Ser Glu Asp Ala Leu Lys
      195              200              205
Leu Ile Arg Ala Ala Thr Gly Lys Ala Ala Ala Ala Leu
      210              215              220

```

<210> 178
 <211> 167
 <212> PRT
 <213> Haemophilus influenzae

```

<400> 178
Met Lys Ile Lys Lys Leu Leu Lys Asn Gly Leu Ser Leu Phe Leu Thr
 1              5              10              15
Phe Ile Val Ile Thr Ser Ile Leu Asp Phe Val Arg Arg Pro Val Val
      20              25              30
Pro Glu Glu Ile Asn Lys Ile Thr Leu Gln Asp Leu Gln Gly Asn Thr
      35              40              45
Phe Ser Leu Glu Ser Leu Asp Gln Asn Lys Pro Thr Leu Leu Tyr Phe
      50              55              60
Trp Gly Thr Trp Cys Gly Tyr Cys Arg Tyr Thr Ser Pro Ala Ile Asn
      65              70              75              80
Ser Leu Ala Lys Glu Gly Tyr Gln Val Val Ser Val Ala Leu Arg Ser
      85              90              95
Gly Asn Glu Ala Asp Val Asn Asp Tyr Leu Ser Lys Asn Asp Tyr His
      100              105              110
Phe Thr Thr Val Asn Asp Pro Lys Gly Glu Phe Ala Glu Arg Trp Gln
      115              120              125
Ile Asn Val Thr Pro Thr Ile Val Leu Leu Ser Lys Gly Lys Met Asp
      130              135              140
Leu Val Thr Thr Gly Leu Thr Ser Tyr Trp Gly Leu Lys Val Arg Leu
      145              150              155              160
Phe Phe Ala Glu Phe Phe Gly
      165

```

<210> 179
 <211> 163
 <212> PRT
 <213> Leishmania major

```

<400> 179
Met Leu Lys Val Ser Ser Lys Glu His Tyr Ala Glu Ile Lys Lys Lys
 1              5              10              15
Ala Glu Asp Ser Leu Gly Leu Val Val His Phe Ser Ala Thr Trp Cys
      20              25              30
Glu Pro Cys Thr Ala Val Asn Glu His Leu Thr Lys Gln Ala Ala Glu
      35              40              45
Tyr Gly Asp Asn Val Val Phe Ala Glu Val Asp Cys Gly Glu Leu Gly
      50              55              60
Asp Val Cys Glu Ala Glu Gly Val Glu Ser Val Pro Phe Val Ala Tyr
      65              70              75              80
Phe Arg Thr Pro Leu Val Gly Asp Asp Arg Arg Val Glu Arg Val Ala
      85              90              95
Asp Val Ala Gly Ala Lys Phe Asp Gln Ile Asp Met Asn Thr His Ser
      100              105              110
Leu Phe Gly Glu Lys Gly Gly Asn Arg Gly Ser Ala Glu Gly Leu Cys
      115              120              125
His Ser Gly Arg Leu Pro Ala Leu Pro His Glu Ala Ala Arg Gly Arg
      130              135              140

```

Asn Val His His Arg His Pro Ile Ser Ser Ala Leu Arg Leu Tyr Trp
 145 150 155 160
 Ser Ala Val

<210> 180
 <211> 275
 <212> PRT
 <213> Mortierella alpina

<400> 180
 Met Val Ser Asn Asn Tyr Ile Asp Ile Thr Ser Glu Asp Asp Phe Ala
 1 5 10 15
 Gln Val Phe Gln Pro Ser Ser Ser Thr Val Tyr Ala Leu Asn Phe Trp
 20 25 30
 Ala Ala Trp Ala Pro Pro Cys Val Gln Met Asn Glu Val Phe Glu Glu
 35 40 45
 Leu Ala Ala Lys Asn Ala Asn Val Asn Phe Leu Lys Ile Glu Ala Glu
 50 55 60
 Lys Phe Pro Asp Ile Ser Glu Asp Tyr Glu Ile Ala Ala Val Pro Ser
 65 70 75 80
 Phe Val Ile Val Lys Glu Gly Thr Val Val Asp Arg Val Glu Gly Ala
 85 90 95
 Asn Ala Pro Glu Leu Ala Lys Val Ile Ala Lys Tyr Ser Lys Ser Thr
 100 105 110
 Ser Ser Pro Leu Pro Thr Gln Ser Thr Met Ala Ala Gly His
 115 120 125
 Ala Ala Pro Ser Val Ala Pro Pro Thr Met Ser Pro Glu Glu Met Asn
 130 135 140
 Ala Arg Leu Lys Glu Leu Thr Ser Ser Ser Ser Val Met Ala Phe Ile
 145 150 155 160
 Lys Gly Thr Pro Thr Ala Pro Arg Cys Gln Phe Ser Arg Gln Leu Leu
 165 170 175
 Glu Ile Leu Thr Ala Gln Asn Ile Arg Phe Ser Ser Phe Asn Ile Leu
 180 185 190
 Ala Asp Asp Glu Val Arg Gln Ala Met Lys Thr Phe Ser Asp Trp Pro
 195 200 205
 Thr Phe Pro Gln Val Tyr Val Lys Gly Glu Phe Val Gly Gly Leu Asp
 210 215 220
 Val Val Lys Glu Leu Val Ala Ser Gly Glu Phe Gln Ala Leu Val Pro
 225 230 235 240
 Ala Glu Lys Asp Leu Lys Thr Arg Met Asp Glu Leu Ile Arg Lys Ala
 245 250 255
 Pro Val Met Ile Phe Ile Lys Gly Ser Pro Glu Thr Pro Arg Cys Gly
 260 265 270
 Phe Ser Lys
 275

<210> 181
 <211> 160
 <212> PRT
 <213> Neisseria gonorrhoeae

<400> 181
 Met Lys Arg Leu Ile Leu Ala Ala Ile Ala Leu Ala Ala Thr Phe Gly
 1 5 10 15
 Ala His Thr Ala Ser Gly Asp Glu Leu Ala Gly Trp Lys Asp Asn Thr
 20 25 30
 Pro Gln Asn Leu Gln Ser Leu Lys Ala Pro Val Arg Ile Ala Asn Leu
 35 40 45
 Trp Ala Thr Trp Cys Gly Pro Cys Arg Lys Glu Met Pro Ala Met Ser
 50 55 60
 Lys Trp Tyr Lys Ala Gln Lys Lys Gly Ser Val Asp Met Val Gly Ile
 65 70 75 80
 Ala Leu Asp Thr Ser Asp Asn Ile Gly Asn Phe Leu Lys Gln Thr Pro

Val	Ser	Tyr	Pro	85	Ile	Trp	Arg	Tyr	Thr	90	Gly	Ala	Asn	Ser	Arg	95	Ser	Phe
			100						105							110		
Met	Lys	Ser	Tyr	Gly	Asn	Asn	Val	Gly	Val	Leu	Pro	Phe	Thr	Val	Val			
		115					120						125					
Glu	Ala	Pro	Lys	Cys	Gly	Tyr	Arg	Gln	Thr	Ile	Thr	Gly	Glu	Leu	Asn			
	130					135						140						
Glu	Lys	Ser	Leu	Thr	Glu	Ala	Val	Lys	Leu	Ala	His	Ser	Lys	Cys	Arg			
145					150					155					160			

<210> 182

<211> 208

<212> PRT

<213> Rhizobium loti

<400> 182

Met	Ala	Gly	Ala	Leu	Ala	Gly	Ala	Val	Ala	Val	Tyr	Val	Ser	Glu	Ser			
1				5				10						15				
Arg	Ser	Gly	Asn	Asn	Ala	Pro	Ala	Arg	Val	Ala	Val	Gly	Gly	Ser	Lys			
		20						25					30					
Asp	Asp	Val	Ala	Cys	Ala	Ala	Lys	Ser	Gly	Arg	Ala	Lys	Lys	Ile	Ala			
		35					40					45						
Ala	Ala	Ala	Thr	Gly	Glu	Val	Ala	Ala	Leu	Leu	Pro	Ala	Asp	Pro	Pro			
	50					55					60							
Gln	Ser	Met	Lys	Ser	Leu	Ala	Phe	Asn	Gly	Pro	Asp	Gly	Lys	Pro	Met			
65					70					75					80			
Thr	Ile	Ala	Asp	His	Ala	Gly	Lys	Thr	Val	Leu	Leu	Asn	Leu	Trp	Ala			
				85					90					95				
Thr	Trp	Cys	Ala	Pro	Cys	Arg	Ala	Glu	Met	Pro	Ala	Leu	Asn	Ala	Leu			
		100						105					110					
Gln	Lys	Asp	Lys	Gly	Ser	Asp	Ala	Phe	Gln	Val	Ile	Ala	Val	Asn	Val			
		115					120					125						
Asp	Ala	Gly	Asp	Asp	Val	Lys	Pro	Lys	Lys	Phe	Leu	Lys	Glu	Thr	Gly			
	130					135					140							
Val	Glu	Ala	Leu	Gly	Tyr	Phe	Arg	Asp	Ser	Thr	Val	Ala	Leu	Phe	Asn			
145					150					155					160			
Asp	Leu	Lys	Ala	Arg	Gly	Leu	Ala	Leu	Gly	Leu	Pro	Val	Thr	Met	Leu			
				165					170					175				
Ile	Asp	Ser	Glu	Gly	Cys	Leu	Ile	Ala	His	Met	Asn	Gly	Pro	Ala	Glu			
		180						185					190					
Trp	Ser	Gly	Arg	Asp	Ala	Arg	Arg	Leu	Val	Glu	Thr	Ala	Leu	Gly	Ser			
		195					200					205						

<210> 183

<211> 176

<212> PRT

<213> Rhodobacter capsulatus

<400> 183

Met	Ala	Lys	Pro	Leu	Met	Phe	Leu	Pro	Leu	Leu	Val	Met	Ala	Gly	Phe			
1				5				10						15				
Val	Gly	Ala	Gly	Tyr	Phe	Ala	Met	Gln	Gln	Asn	Asp	Pro	Asn	Ala	Met			
		20						25					30					
Pro	Thr	Ala	Leu	Ala	Gly	Lys	Glu	Ala	Pro	Ala	Val	Arg	Leu	Glu	Pro			
		35					40					45						
Leu	Gly	Ala	Glu	Ala	Pro	Phe	Thr	Asp	Ala	Asp	Leu	Arg	Asp	Gly	Lys			
	50					55					60							
Ile	Lys	Leu	Val	Asn	Phe	Trp	Ala	Ser	Trp	Cys	Ala	Pro	Cys	Arg	Val			
65				70						75				80				
Glu	His	Pro	Asn	Leu	Ile	Gly	Leu	Lys	Gln	Asp	Gly	Ile	Glu	Ile	Met			
				85					90					95				
Gly	Val	Asn	Trp	Lys	Asp	Thr	Pro	Asp	Gln	Ala	Gln	Gly	Phe	Leu	Ala			
		100					105						110					
Glu	Met	Gly	Ser	Pro	Tyr	Thr	Arg	Leu	Gly	Ala	Asp	Pro	Gly	Asn	Lys			
		115					120					125						

Met Gly Leu Asp Trp Gly Val Ala Gly Val Pro Glu Thr Phe Val Val
 130 135 140
 Asp Gly Ala Gly Arg Ile Leu Thr Arg Ile Ala Gly Pro Leu Thr Glu
 145 150 155 160
 Asp Val Ile Thr Lys Lys Ile Asp Pro Leu Leu Ala Gly Thr Ala Asp
 165 170 175

<210> 184
 <211> 105
 <212> PRT
 <213> Synechocystis

<400> 184
 Met Ala Val Lys Lys Gln Phe Ala Asn Phe Ala Glu Met Leu Ala Gly
 1 5 10 15
 Ser Pro Lys Pro Val Leu Val Asp Phe Tyr Ala Thr Trp Cys Gly Pro
 20 25 30
 Cys Gln Met Met Ala Pro Ile Leu Glu Gln Val Gly Ser His Leu Arg
 35 40 45
 Gln Gln Ile Gln Val Val Lys Ile Asp Thr Asp Lys Tyr Pro Ala Ile
 50 55 60
 Ala Thr Gln Tyr Gln Ile Gln Ser Leu Pro Thr Leu Val Leu Phe Lys
 65 70 75 80
 Gln Gly Gln Pro Val His Arg Met Glu Gly Val Gln Gln Ala Ala Gln
 85 90 95
 Leu Ile Gln Gln Leu Gln Val Phe Val
 100 105

<210> 185
 <211> 109
 <212> PRT
 <213> Synechocystis

<400> 185
 Met Ser Leu Leu Glu Ile Thr Asp Ala Glu Phe Glu Gln Glu Thr Gln
 1 5 10 15
 Gly Gln Thr Lys Pro Val Leu Val Tyr Phe Trp Ala Ser Trp Cys Gly
 20 25 30
 Pro Cys Arg Leu Met Ala Pro Ala Ile Gln Ala Ile Lys Asp Tyr
 35 40 45
 Gly Asp Lys Leu Lys Val Leu Lys Leu Glu Val Asp Pro Asn Pro Ala
 50 55 60
 Ala Val Ala Gln Cys Lys Val Glu Gly Val Pro Ala Leu Arg Leu Phe
 65 70 75 80
 Lys Asn Asn Glu Leu Val Met Thr His Glu Gly Ala Ile Ala Lys Pro
 85 90 95
 Lys Leu Leu Glu Leu Leu Lys Glu Glu Leu Asp Phe Ile
 100 105

<210> 186
 <211> 290
 <212> PRT
 <213> Schizosaccharomyces pombe

<400> 186
 Met Ser Val Ile Glu Ile Arg Ser Tyr Gln His Trp Ile Ser Thr Ile
 1 5 10 15
 Pro Lys Ser Gly Tyr Leu Ala Val Asp Cys Tyr Ala Asp Trp Cys Gly
 20 25 30
 Pro Cys Lys Ala Ile Ser Pro Leu Phe Ser Gln Leu Ala Ser Lys Tyr
 35 40 45
 Ala Ser Pro Lys Phe Val Phe Ala Lys Val Asn Val Asp Glu Gln Arg
 50 55 60
 Gln Ile Ala Ser Gly Leu Gly Val Lys Ala Met Pro Thr Phe Val Phe

65	Phe	Glu	Asn	Gly	Lys	Gln	Ile	Asp	Met	Leu	Thr	Gly	Ala	Asn	Pro	Gln
					85					90					95	
	Ala	Leu	Lys	Glu	Lys	Val	Ala	Leu	Ile	Ser	Ser	Lys	Ala	Thr	Gly	Thr
				100					105					110		
	Gly	Ala	Leu	Ala	Ser	Ser	Ser	Ser	Ala	Pro	Val	Lys	Gly	Phe	Ala	Ser
			115					120					125			
	Leu	Gln	Gly	Cys	Ile	Glu	Asn	Pro	Gln	Leu	Glu	Cys	Leu	Asn	Gln	Gln
			130				135					140				
	Asp	Asp	His	Asp	Leu	Lys	Ser	Ala	Phe	Asn	Ser	Asn	Pro	Ser	Ser	Phe
	145					150					155					160
	Leu	Glu	Ser	Asp	Val	Asp	Glu	Gln	Leu	Met	Ile	Tyr	Ile	Pro	Phe	Leu
				165						170					175	
	Glu	Val	Val	Lys	Val	His	Ser	Ile	Ala	Ile	Thr	Pro	Val	Lys	Gly	Glu
				180					185					190		
	Thr	Ser	Ser	Ala	Pro	Lys	Thr	Ile	Lys	Leu	Tyr	Ile	Asn	Gln	Pro	Asn
				195				200					205			
	Asn	Leu	Ser	Phe	Glu	Asp	Ala	Glu	Ser	Phe	Thr	Pro	Thr	Gln	Val	Ile
				210			215					220				
	Glu	Asp	Ile	Val	Tyr	Glu	Gln	Asp	Asp	Gln	Pro	Thr	Ile	Ile	Pro	Leu
	225					230					235					240
	Arg	Phe	Val	Lys	Phe	Gln	Arg	Val	Asn	Ser	Leu	Val	Ile	Phe	Ile	Tyr
				245						250					255	
	Ser	Asn	Val	Gly	Glu	Glu	Glu	Thr	Thr	Lys	Ile	Ser	Arg	Leu	Glu	Leu
				260					265					270		
	Phe	Gly	Glu	Pro	Val	Gly	Asp	Ser	Ser	Lys	Gly	Lys	Leu	Gln	Lys	Val
			275					280					285			
	Glu	Ala														
		290														

<210> 187

<211> 185

<212> PRT

<213> Treponema pallidum

<400> 187

Met	Phe	Arg	Ser	Asp	Leu	Val	Leu	Ala	Val	Trp	Gly	Val	Thr	Cys	Val
1				5					10					15	
Gln	Ala	Ala	Asp	Val	Ala	His	Asn	Ala	Asp	Val	Pro	Ser	Arg	Ser	Leu
			20					25					30		
Lys	Ala	Leu	Glu	Arg	Phe	Arg	Phe	Phe	Val	Tyr	Pro	Lys	Pro	Leu	Asp
		35					40					45			
Leu	Ser	Ser	Asp	Phe	His	Ala	Lys	Ala	Leu	Lys	Gly	Glu	Ala	Leu	Val
	50				55						60				
Pro	Ser	Leu	Phe	Lys	Gly	Lys	Val	Thr	Leu	Leu	Asn	Phe	Trp	Ala	Thr
65					70				75						80
Trp	Cys	Pro	Pro	Cys	Arg	Ala	Glu	Met	Pro	Ser	Met	Asp	Arg	Met	Gln
				85				90						95	
Ala	Leu	Met	Arg	Gly	Asn	Asp	Phe	Gln	Ile	Val	Ala	Val	Asn	Val	Gly
			100					105					110		
Asp	Ser	Arg	Lys	Gln	Val	Glu	Ser	Phe	Ile	Ala	Arg	Gly	Lys	His	Thr
		115					120					125			
Phe	Pro	Ile	Tyr	Leu	Asp	Glu	Glu	Gly	Ser	Leu	Gly	Ser	Val	Phe	Ala
	130					135					140				
Ser	Arg	Gly	Leu	Pro	Thr	Tyr	Val	Val	Asp	Lys	Ala	Gly	Arg	Ile	
145					150				155					160	
Val	Ala	Val	Val	Val	Gly	Ser	Val	Glu	Tyr	Asp	Gln	Pro	Glu	Leu	Val
				165				170						175	
Ala	Leu	Phe	Lys	Glu	Leu	Ala	Arg	Asp							
			180					185							

<210> 188

<211> 246

<212> PRT

<213> Caenorhabditis elegans

<400> 188

```

Met Leu Leu Arg Leu Leu Ala Val Leu Gly Leu Phe Ala Val Gly Val
1      5      10      15
Ser Gly Gly Pro Thr Arg Ser Ser Lys Leu Val Phe Leu Asn Glu Glu
20     25     30
Asn Trp Thr Asp Leu Met Lys Gly Glu Trp Met Ile Glu Phe His Ala
35     40     45
Pro Trp Cys Pro Ala Cys Lys Asp Leu Gln Lys Ala Trp Asn Ala Phe
50     55     60
Ala Asp Trp Ser Asp Asp Leu Gly Ile Lys Val Gly Glu Val Asp Val
65     70     75     80
Thr Val Asn Pro Gly Leu Ser Gly Arg Phe Leu Val Thr Ala Leu Pro
85     90     95
Thr Ile Tyr His Val Lys Asp Gly Val Phe Arg Gln Tyr Ser Gly Ala
100    105    110
Arg Asp Lys Asn Asp Phe Ile Ser Phe Val Glu Asp Lys Lys Tyr Arg
115    120    125
Val Ile Asp Pro Val Pro Asp Tyr Lys His Pro Asn Ser Lys Gln Met
130    135    140
Ala Val Val Ala Val Phe Phe Lys Leu Ser Met Ser Val Arg Asp Leu
145    150    155    160
His Asn His Leu Val Glu Asp Lys Gly Ile Pro Ser Trp Ala Ser Tyr
165    170    175
Gly Leu Phe Ala Gly Val Thr Leu Ala Leu Gly Cys Val Leu Gly Phe
180    185    190
Phe Ile Val Ile Ile Ile Asp Gln Val Phe Pro Thr Gly Pro Arg Lys
195    200    205
Ser Gln Gln Ala Lys Lys Thr Glu Lys Lys Asp Ala Lys Lys Asp Ser
210    215    220
Gly Thr Glu Ser Pro Thr Lys Lys Asn Gly Asn Asn Asn Asn Gly Lys
225    230    235    240
Glu Thr Lys Lys Thr Lys
245

```

<210> 189

<211> 284

<212> PRT

<213> Caenorhabditis elegans

<400> 189

```

Met Pro Val Ile Asn Val Lys Asp Asp Glu Asp Phe Arg Asn Gln Leu
1      5      10      15
Ser Leu Ala Gly Leu Lys Ser Val Ile Val Asp Phe Thr Ala Val Trp
20     25     30
Cys Gly Pro Cys Lys Met Ile Ala Pro Thr Phe Glu Ala Leu Ser Asn
35     40     45
Gln Tyr Leu Gly Ala Val Phe Leu Lys Val Asp Val Glu Ile Cys Glu
50     55     60
Lys Thr Ser Ser Glu Asn Gly Val Asn Ser Met Pro Thr Phe Met Val
65     70     75     80
Phe Gln Ser Gly Val Arg Val Glu Gln Met Lys Gly Ala Asp Ala Lys
85     90     95
Ala Leu Glu Thr Met Val Lys Lys Tyr Ala Asp Asn Ser Ala Ala Asp
100    105    110
Ser Leu Val Ala Gly Gln Met Asp Leu Thr Pro Leu Val Asp Lys Lys
115    120    125
Gln Met Glu Cys Leu Asn Glu Ser Asp Asp Thr Pro Leu Gly Arg Phe
130    135    140
Leu Glu Gly Asn Cys Asn Leu Val Ser Asp Cys Asp Glu Gln Leu Ile
145    150    155    160
Ile Ser Leu Pro Phe Asn Gln Pro Val Lys Val His Ser Ile Leu Ile
165    170    175
Lys Gly Val Ser Asp Arg Ala Pro Lys Lys Val Lys Val Phe Ile Asn
180    185    190
Leu Pro Lys Thr Thr Asp Phe Asp Asn Ala Thr Ala Leu Glu Pro Thr
195    200    205

```

Gln Met Leu Glu Phe Asp Glu Ser Ser Ile Gln Gly His Gly Gln Val
 210 215 220
 Val Ala Leu Lys Tyr Val Lys Phe Gln Asn Val Gln Asn Ile Gln Phe
 225 230 235 240
 Phe Ile Glu Asn Asn Val Gly Gly Gly Asp Val Thr Glu Leu Val Lys
 245 250 255
 Leu Thr Val Phe Gly Thr Pro Leu Ser Ala Leu Asn Met Asn Glu Phe
 260 265 270
 Lys Arg Val Ala Gly Lys Ala Gly Asp Ala Ala His
 275 280

<210> 190

<211> 287

<212> PRT

<213> *Drosophila melanogaster*

<400> 190

Met Ser Val Arg Val Ile Asn Asp Glu Ser His Phe Gln Ala Glu Leu
 1 5 10 15
 Ala Gln Ala Gly Ile Gln Leu Val Val Asp Phe Thr Ala Ser Trp
 20 25 30
 Cys Gly Pro Cys Lys Arg Ile Ala Pro Ile Phe Glu Thr Phe Pro Thr
 35 40 45
 Lys Tyr Pro Lys Ala Ile Phe Leu Lys Val Asp Val Asp Lys Cys Gln
 50 55 60
 Asp Thr Ala Ala Gly Gln Gly Val Ser Ala Met Pro Thr Phe Ile Phe
 65 70 75 80
 Tyr Arg Asn Arg Thr Lys Ile Asp Arg Val Gln Gly Ala Asp Val Asn
 85 90 95
 Gly Leu Glu Ala Lys Ile Gln Glu His Ile Gly Thr Ser Gly Gly Glu
 100 105 110
 Glu Gly Gly Glu Asp Tyr Gly Gln Gly Leu Met Glu Leu Asn Thr Phe
 115 120 125
 Ile Ser Lys Gln Glu Cys Glu Cys Leu Asn Glu Ala Asp Asp His Asn
 130 135 140
 Leu Lys His Ala Leu Ala Ser Ala Gly Gly Tyr Leu Gln Ser Asp Cys
 145 150 155 160
 Asp Glu Gln Leu Ile Leu Ser Ile Thr Phe Asn Gln Ala Val Lys Ile
 165 170 175
 His Ser Leu Lys Phe Lys Ala Pro Ser His Leu Gly Pro Lys Asp Val
 180 185 190
 Lys Leu Phe Ile Asn Gln Pro Arg Thr Ile Asp Phe Asp Met Ala Glu
 195 200 205
 Ser Met Asn Ser Val Gln Asp Leu Ser Leu Ala Gln Lys Glu Leu Glu
 210 215 220
 Ser Gly Val Pro Val Asn Leu Arg Tyr Val Lys Phe Gln Asn Val Gln
 225 230 235 240
 Asn Ile Gln Ile Phe Val Lys Asn Asn Gln Ser Gly Gly Asp Val Thr
 245 250 255
 Gln Ile Asp Tyr Ile Gly Phe Ile Gly Ser Pro Ile Met Thr Thr Lys
 260 265 270
 Met Asn Asp Phe Lys Arg Val Ala Gly Lys Lys Gly Glu Ser His
 275 280 285

<210> 191

<211> 289

<212> PRT

<213> *Homo sapien*

<400> 191

Met Val Gly Val Lys Pro Val Gly Ser Asp Pro Asp Phe Gln Pro Glu
 1 5 10 15
 Leu Ser Gly Ala Gly Ser Arg Leu Ala Val Val Lys Phe Thr Met Arg
 20 25 30
 Gly Cys Gly Pro Cys Leu Arg Ile Ala Pro Ala Phe Ser Ser Met Ser

```
<210> 192
<211> 335
<212> PRT
<213> Homo sapien
```

-106-

```

Leu Ile Gly Gly Leu Asp Ile Ile Lys Glu Leu Glu Ala Ser Glu Glu
 210                215                220
Leu Asp Thr Ile Cys Pro Lys Ala Pro Lys Leu Glu Glu Arg Leu Lys
225                230                235                240
Val Leu Thr Asn Lys Ala Ser Val Met Leu Phe Met Lys Gly Asn Lys
                245                250                255
Gln Glu Ala Lys Cys Gly Phe Ser Lys Gln Ile Leu Glu Ile Leu Asn
                260                265                270
Ser Thr Gly Val Glu Tyr Glu Thr Phe Asp Ile Leu Glu Asp Glu Glu
                275                280                285
Val Arg Gln Gly Leu Lys Ala Tyr Ser Asn Trp Pro Thr Tyr Pro Gln
                290                295                300
Leu Tyr Val Lys Gly Glu Leu Val Gly Gly Leu Asp Ile Val Lys Glu
305                310                315                320
Leu Lys Glu Asn Gly Glu Leu Leu Pro Ile Leu Arg Gly Glu Asn
                325                330                335

```

<210> 193
 <211> 131
 <212> PRT
 <213> *Phalaris coerulescens*

```

<400> 193
Met Gly Gly Cys Val Gly Lys Asp Arg Gly Ile Val Glu Asp Lys Leu
 1                5                10                15
Asp Phe Lys Gly Asn Val His Val Ile Thr Thr Lys Glu Asp Trp
                20                25                30
Asp Gln Lys Ile Ala Glu Ala Asn Lys Asp Gly Lys Ile Val Val Ala
                35                40                45
Asn Phe Ser Ala Ser Trp Cys Gly Pro Cys Arg Val Ile Ala Pro Val
                50                55                60
Tyr Ala Glu Met Ser Lys Thr Tyr Pro Gln Leu Met Phe Leu Thr Ile
65                70                75                80
Asp Val Asp Asp Leu Val Asp Phe Ser Ser Thr Trp Asp Ile Arg Ala
                85                90                95
Thr Pro Thr Phe Phe Leu Lys Asn Gly Gln Gln Ile Asp Lys Leu
                100                105                110
Val Gly Ala Asn Lys Pro Glu Leu Glu Lys Lys Val Gln Ala Leu Gly
                115                120                125
Asp Gly Ser
130

```

<210> 194
 <211> 144
 <212> PRT
 <213> *Trypanosoma brucei brucei*

```

<400> 194
Met Ser Gly Leu Ala Lys Tyr Leu Pro Gly Ala Thr Asn Leu Leu Ser
 1                5                10                15
Lys Ser Gly Glu Val Ser Leu Gly Ser Leu Val Gly Lys Thr Val Phe
                20                25                30
Leu Tyr Phe Ser Ala Ser Trp Cys Pro Pro Cys Arg Gly Phe Thr Pro
                35                40                45
Val Leu Ala Glu Phe Tyr Glu Lys His His Val Ala Lys Asn Phe Glu
                50                55                60
Val Val Leu Ile Ser Trp Asp Glu Asn Glu Ser Asp Phe His Asp Tyr
65                70                75                80
Tyr Gly Lys Met Pro Trp Leu Ala Leu Pro Phe Asp Gln Arg Ser Thr
                85                90                95
Val Ser Glu Leu Gly Lys Thr Phe Gly Val Glu Ser Ile Pro Thr Leu
                100                105                110
Ile Thr Ile Asn Ala Asp Thr Gly Ala Ile Ile Gly Thr Gln Ala Arg
                115                120                125
Thr Arg Val Ile Glu Asp Pro Asp Gly Ala Asn Phe Pro Trp Pro Asn

```

130

135

140

<210> 195
 <211> 333
 <212> PRT
 <213> Arabidopsis thaliana

<400> 195
 Met Asn Gly Leu Glu Thr His Asn Thr Arg Leu Cys Ile Val Gly Ser
 1 5 10 15
 Gly Pro Ala Ala His Thr Ala Ala Ile Tyr Ala Ala Arg Ala Glu Leu
 20 25 30
 Lys Pro Leu Leu Phe Glu Gly Trp Met Ala Asn Asp Ile Ala Pro Gly
 35 40 45
 Gly Gln Leu Thr Thr Thr Thr Asp Val Glu Asn Phe Pro Gly Phe Pro
 50 55 60
 Glu Gly Ile Leu Gly Val Glu Leu Thr Asp Lys Phe Arg Lys Gln Ser
 65 70 75 80
 Glu Arg Phe Gly Thr Thr Ile Phe Thr Glu Thr Val Thr Lys Val Asp
 85 90 95
 Phe Ser Ser Lys Pro Phe Lys Leu Phe Thr Asp Ser Lys Ala Ile Leu
 100 105 110
 Ala Asp Ala Val Ile Leu Ala Thr Gly Ala Val Ala Lys Arg Leu Ser
 115 120 125
 Phe Val Gly Ser Gly Glu Ala Ser Gly Gly Phe Trp Asn Arg Gly Ile
 130 135 140
 Ser Ala Cys Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg Asn Lys
 145 150 155 160
 Pro Leu Ala Val Ile Gly Gly Gly Asp Ser Ala Met Glu Glu Ala Asn
 165 170 175
 Phe Leu Thr Lys Tyr Gly Ser Lys Val Tyr Ile Ile His Arg Arg Asp
 180 185 190
 Ala Phe Arg Ala Ser Lys Ile Met Gln Gln Arg Ala Leu Ser Asn Pro
 195 200 205
 Lys Ile Asp Val Ile Trp Asn Ser Ser Val Val Glu Ala Tyr Gly Asp
 210 215 220
 Gly Glu Arg Asp Val Leu Gly Gly Leu Lys Val Lys Asn Val Val Thr
 225 230 235 240
 Gly Asp Val Ser Asp Leu Lys Val Ser Gly Leu Phe Phe Ala Ile Gly
 245 250 255
 His Glu Pro Ala Thr Lys Phe Leu Asp Gly Gly Val Glu Leu Asp Ser
 260 265 270
 Asp Gly Tyr Val Val Thr Lys Pro Gly Thr Thr Gln Thr Ser Val Pro
 275 280 285
 Gly Val Phe Ala Ala Gly Asp Val Gln Asp Lys Lys Tyr Arg Gln Ala
 290 295 300
 Ile Thr Ala Ala Gly Thr Gly Cys Met Ala Ala Leu Asp Ala Glu His
 305 310 315 320
 Tyr Leu Gln Glu Ile Gly Ser Gln Gln Gly Lys Ser Asp
 325 330

<210> 196
 <211> 383
 <212> PRT
 <213> Arabidopsis thaliana

<400> 196
 Met Cys Trp Ile Ser Met Ser Gln Ser Arg Phe Ile Ile Lys Ser Leu
 1 5 10 15
 Phe Ser Thr Ala Gly Gly Phe Leu Leu Gly Ser Ala Leu Ser Asn Pro
 20 25 30
 Pro Ser Leu Ala Thr Ala Phe Ser Ser Ser Ser Ser Ser Ser Ala
 35 40 45
 Ala Ala Ala Val Asp Met Glu Thr His Lys Thr Lys Val Cys Ile Val
 50 55 60

Gly Ser Gly Pro Ala Ala His Thr Ala Ala Ile Tyr Ala Ser Arg Ala
 65 70 75 80
 Glu Leu Lys Pro Leu Leu Phe Glu Gly Trp Met Ala Asn Asp Ile Ala
 85 90 95
 Pro Gly Gly Gln Leu Thr Thr Thr Thr Asp Val Glu Asn Phe Pro Gly
 100 105 110
 Phe Pro Glu Gly Ile Leu Gly Ile Asp Ile Val Glu Lys Phe Arg Lys
 115 120 125
 Gln Ser Glu Arg Phe Gly Thr Thr Ile Phe Thr Glu Thr Val Asn Lys
 130 135 140
 Val Asp Phe Ser Ser Lys Pro Phe Lys Leu Phe Thr Asp Ser Arg Thr
 145 150 155 160
 Val Leu Ala Asp Ser Val Ile Ile Ser Thr Gly Ala Val Ala Lys Arg
 165 170 175
 Leu Ser Phe Thr Gly Ser Gly Glu Gly Asn Gly Gly Phe Trp Asn Arg
 180 185 190
 Gly Ile Ser Ala Cys Ala Val Cys Asp Gly Ala Ala Pro Ile Phe Arg
 195 200 205
 Asn Lys Pro Leu Val Val Ile Gly Gly Gly Asp Ser Ala Met Glu Glu
 210 215 220
 Ala Asn Phe Leu Thr Lys Tyr Gly Ser Lys Val Tyr Ile Ile His Arg
 225 230 235 240
 Arg Asp Thr Phe Arg Ala Ser Lys Ile Met Gln Gln Arg Ala Leu Ser
 245 250 255
 Asn Pro Lys Ile Glu Val Ile Trp Asn Ser Ala Val Val Glu Ala Tyr
 260 265 270
 Gly Asp Glu Asn Gly Arg Val Leu Gly Gly Leu Lys Val Lys Asn Val
 275 280 285
 Val Thr Gly Asp Val Ser Asp Leu Lys Val Ser Gly Leu Phe Phe Ala
 290 295 300
 Ile Gly His Glu Pro Ala Thr Lys Phe Leu Asp Gly Gln Leu Glu Leu
 305 310 315 320
 Asp Glu Asp Gly Tyr Val Val Thr Lys Pro Gly Thr Thr Lys Thr Ser
 325 330 335
 Val Val Gly Val Phe Ala Ala Gly Asp Val Gln Asp Lys Lys Tyr Arg
 340 345 350
 Gln Ala Ile Thr Ala Ala Gly Thr Gly Cys Met Ala Ala Leu Asp Ala
 355 360 365
 Glu His Tyr Leu Gln Glu Ile Gly Ser Gln Glu Gly Lys Ser Asp
 370 375 380

<210> 197

<211> 323

<212> PRT

<213> Aquifex aeolicus

<400> 197

Met Ala Val Ser Leu Met Gln Gln Pro Asp Lys Val Tyr Asp Val Ile
 1 5 10 15
 Ile Ile Gly Ala Gly Pro Ala Gly Thr Thr Ala Ala Ile Tyr Thr Ala
 20 25 30
 Arg Ala Gly Trp Lys Thr Leu Val Leu Tyr Arg Ala Glu Ala Asp Gly
 35 40 45
 Ala Leu Gly Val Thr Gln Lys Ile Glu Asn Tyr Pro Gly Val Pro Gly
 50 55 60
 Pro Leu Ser Gly Tyr Glu Leu Leu Lys Ile Met Arg Glu Gln Ala Lys
 65 70 75 80
 Ser Phe Gly Ala Glu Phe Val Arg Gly Lys Val Ile Ala Thr Asp Leu
 85 90 95
 Asn Ser Asp Pro Lys Lys Val Tyr Thr Ile Asp Gly Arg Glu Phe Arg
 100 105 110
 Gly Lys Thr Ile Ile Val Ala Ser Gly Ala Met Glu Arg Ala Asn Lys
 115 120 125
 Phe Lys Gly Glu Glu Glu Phe Leu Gly Arg Gly Val Ser Tyr Cys Gly
 130 135 140
 Val Cys Asp Ala Ala Phe Phe Lys Asp Gln Pro Val Ala Val Ile Gly

145	Asp	Asp	Asp	Tyr	Ala	Ile	Glu	Glu	Ala	Glu	Phe	Ile	Ala	Arg	Phe	Ala
					165	Val	Val	Pro	Gly	Ser	Lys	Ile	Lys	Ala	Pro	Pro
					180	His	Phe	Glu	Lys	Leu	Pro	Asn	Val	Glu	Ile	Leu
					195											
					210											
					225											
					245											
					260											
					275											
					290											
					305											
					310											

<210> 198

<211> 315

<212> PRT

<213> Bacillus subtilis

<400> 198

Ser	Glu	Glu	Lys	Ile	Tyr	Asp	Val	Ile	Ile	Ile	Gly	Ala	Gly	Pro	Ala
1				5				10						15	
Gly	Met	Thr	Ala	Ala	Val	Tyr	Thr	Ser	Arg	Ala	Asn	Leu	Ser	Thr	Leu
				20				25					30		
Met	Ile	Glu	Arg	Gly	Ile	Pro	Gly	Gly	Gln	Met	Ala	Asn	Thr	Glu	Asp
				35			40					45			
Val	Glu	Asn	Tyr	Pro	Gly	Phe	Glu	Ser	Ile	Leu	Gly	Pro	Glu	Leu	Ser
				50		55					60				
Asn	Lys	Met	Phe	Glu	His	Ala	Lys	Lys	Phe	Gly	Ala	Glu	Tyr	Ala	Tyr
65				70					75					80	
Gly	Asp	Ile	Lys	Glu	Val	Ile	Asp	Gly	Lys	Glu	Tyr	Lys	Val	Val	Lys
				85				90					95		
Ala	Gly	Ser	Lys	Glu	Tyr	Lys	Ala	Arg	Ala	Val	Ile	Ile	Ala	Ala	Gly
				100				105					110		
Ala	Glu	Tyr	Lys	Lys	Ile	Gly	Val	Pro	Gly	Glu	Lys	Glu	Leu	Gly	Gly
				115			120					125			
Arg	Gly	Val	Ser	Tyr	Cys	Ala	Val	Cys	Asp	Gly	Ala	Phe	Phe	Lys	Gly
130					135						140				
Lys	Glu	Leu	Val	Val	Val	Gly	Gly	Gly	Asp	Ser	Ala	Val	Glu	Glu	Gly
145					150				155					160	
Val	Tyr	Leu	Thr	Arg	Phe	Ala	Ser	Lys	Val	Thr	Ile	Val	His	Arg	Arg
				165					170					175	
Asp	Lys	Leu	Arg	Ala	Gln	Ser	Ile	Leu	Gln	Ala	Arg	Ala	Phe	Asp	Asn
				180				185					190		
Glu	Lys	Val	Asp	Phe	Leu	Trp	Asn	Lys	Thr	Val	Lys	Glu	Ile	His	Glu
				195			200					205			
Glu	Asn	Gly	Lys	Val	Gly	Asn	Val	Thr	Leu	Val	Asp	Thr	Val	Thr	Gly
210					215						220				
Glu	Glu	Ser	Glu	Phe	Lys	Thr	Asp	Gly	Val	Phe	Ile	Tyr	Ile	Gly	Met
225					230					235				240	
Leu	Pro	Leu	Ser	Lys	Pro	Phe	Glu	Asn	Leu	Gly	Ile	Thr	Asn	Glu	Glu
				245					250					255	
Gly	Tyr	Ile	Glu	Thr	Asn	Asp	Arg	Met	Glu	Thr	Lys	Val	Glu	Gly	Ile
				260				265					270		
Phe	Ala	Ala	Gly	Asp	Ile	Arg	Glu	Lys	Ser	Leu	Arg	Gln	Ile	Val	Thr
				275			280					285			

Ala Thr Gly Asp Gly Ser Ile Ala Ala Gln Ser Val Gln His Tyr Val
 290 295 300
 Glu Glu Leu Gln Glu Thr Leu Lys Thr Leu Lys
 305 310 315

<210> 199

<211> 326

<212> PRT

<213> *Borrelia burgdorferi*

<400> 199

Met Leu Glu Phe Glu Thr Ile Asp Ile Asn Leu Thr Lys Lys Lys Asn
 1 5 10 15
 Leu Ser Gln Lys Glu Val Asp Phe Ile Glu Asp Val Ile Ile Val Gly
 20 25 30
 Ser Gly Pro Ala Gly Leu Thr Ala Gly Ile Tyr Ser Val Met Ser Asn
 35 40 45
 Tyr Lys Ala Ala Ile Leu Glu Gly Pro Glu Pro Gly Gly Gln Leu Thr
 50 55 60
 Thr Thr Thr Glu Val Tyr Asn Tyr Pro Gly Phe Lys Asn Gly Ile Ser
 65 70 75 80
 Gly Arg Asn Leu Met Leu Asn Met Arg Glu Gln Val Val Asn Leu Gly
 85 90 95
 Ala Lys Thr Phe Pro Glu Thr Val Phe Ser Ile Lys Arg Lys Gly Asn
 100 105 110
 Ile Phe Tyr Leu Tyr Thr Glu Asn Tyr Ile Tyr Lys Ser Lys Ala Val
 115 120 125
 Ile Ile Ala Val Gly Ser Lys Pro Lys Lys Leu Glu Thr Leu Lys Asn
 130 135 140
 Ser Gly Leu Phe Trp Asn Lys Gly Ile Ser Val Cys Ala Ile Cys Asp
 145 150 155 160
 Gly His Leu Phe Lys Gly Lys Arg Val Ala Val Ile Gly Gly Gly Asn
 165 170 175
 Thr Ala Leu Ser Glu Ser Ile Tyr Leu Ser Lys Leu Val Asp Lys Val
 180 185 190
 Tyr Leu Ile Val Arg Lys Asn Asn Leu Arg Ala Ile Ala Met Leu Arg
 195 200 205
 Asp Ser Val Ala Lys Leu Pro Asn Ile Glu Ile Leu Tyr Asn Ser Glu
 210 215 220
 Ala Ile Glu Val Asp Gly Lys Ser Ser Val Ser Ser Val Lys Ile Phe
 225 230 235 240
 Asn Lys Lys Asp Asn Val Val Tyr Glu Leu Glu Val Ser Ala Val Phe
 245 250 255
 Met Ala Val Gly Tyr Lys Pro Asn Thr Glu Phe Leu Lys Gly Phe Leu
 260 265 270
 Asp Leu Asp Glu Glu Gly Phe Ile Val Thr Lys Asp Val Val Lys Thr
 275 280 285
 Ser Val Asp Gly Val Phe Ser Cys Gly Asp Val Ser Asn Lys Leu Tyr
 290 295 300
 Ala Gln Ala Ile Thr Ala Ala Ala Glu Gly Phe Ile Ala Ser Val Glu
 305 310 315 320
 Leu Gly Asn Phe Leu Lys
 325

<210> 200

<211> 319

<212> PRT

<213> *Buchnera aphidicola*

<400> 200

Met Asp Lys Val Lys His Ser Lys Ile Ile Ile Leu Gly Ser Gly Pro
 1 5 10 15
 Ala Gly Tyr Thr Ala Ala Ile Tyr Ala Ala Arg Ala Asn Leu Asp Pro
 20 25 30
 Phe Leu Ile Thr Gly Thr Asn Lys Gly Gly Gln Leu Met Asn Thr Asn

		35					40					45			
Glu	Ile	Glu	Asn	Trp	Pro	Gly	Asp	Tyr	Asn	Lys	Ile	Ser	Gly	Ser	Glu
	50					55					60				
Leu	Met	Asn	Arg	Met	Tyr	Lys	His	Ala	Ile	Glu	Leu	Lys	Thr	Lys	Val
65					70					75					80
Ile	Cys	Asp	Thr	Val	Ile	Ser	Val	Asn	Phe	Lys	Lys	Asn	Pro	Phe	Phe
				85					90					95	
Leu	Ile	Gly	Glu	Asn	Asn	Lys	Tyr	Thr	Ala	Asp	Ser	Val	Ile	Ile	Ala
			100					105					110		
Thr	Gly	Ala	Asn	Pro	Arg	Tyr	Leu	Gly	Leu	Gln	Ser	Glu	Ser	Leu	Phe
		115					120					125			
Lys	Gly	Lys	Gly	Val	Ser	Thr	Cys	Ala	Val	Cys	Asp	Gly	Phe	Phe	Tyr
	130					135					140				
Lys	Asn	Lys	Glu	Val	Ala	Val	Val	Gly	Gly	Gly	Asn	Thr	Ala	Ile	Glu
145					150					155					160
Glu	Thr	Leu	Tyr	Leu	Ser	Asn	Phe	Val	Lys	Lys	Val	His	Leu	Ile	His
				165					170					175	
Arg	Gly	Ile	Asn	Phe	Arg	Ala	Glu	Lys	Ile	Leu	Leu	Asp	Arg	Leu	Glu
			180					185					190		
Lys	Lys	Ile	Lys	Ser	Gln	Lys	Ile	Ile	Ile	Tyr	Leu	Asn	Ser	Ile	Val
		195					200					205			
Lys	Asn	Ile	Leu	Gly	Asn	Ser	Ser	Gly	Val	Thr	Ala	Leu	Leu	Ile	Glu
	210					215					220				
Gln	Lys	Asn	Ser	Lys	Glu	Lys	Thr	Glu	Ser	Lys	Ile	Gln	Val	Ser	Gly
225					230					235					240
Leu	Phe	Val	Ala	Ile	Gly	Tyr	Thr	Pro	Asn	Thr	Asn	Ile	Phe	Val	Asn
				245					250					255	
Lys	Leu	Lys	Met	Lys	Asp	Gly	Tyr	Ile	Gln	Val	Thr	Arg	Gln	Glu	His
			260					265					270		
Gly	Asn	Tyr	Thr	Gln	Thr	Ser	Ile	Pro	Gly	Ile	Phe	Ala	Ala	Gly	Asp
		275					280					285			
Val	Ile	Asp	His	Val	Tyr	Arg	Gln	Ala	Ile	Thr	Ser	Ser	Ala	Ser	Gly
	290					295					300				
Cys	Met	Ala	Ala	Leu	Asp	Ser	Glu	Arg	Tyr	Ile	Asn	Ser	Leu	Val	
305					310					315					

```
<210> 201
<211> 319
<212> PRT
<213> Buchnera aphidicola
```

<400>	201															
Met	Glu	Leu	Lys	Asn	His	Lys	Lys	Ile	Ile	Ile	Leu	Gly	Ser	Gly	Pro	
1				5					10					15		
Ala	Gly	Tyr	Thr	Ala	Ala	Ile	Tyr	Ser	Ser	Arg	Ala	Asn	Leu	Asn	Pro	
			20					25					30			
Leu	Leu	Ile	Thr	Gly	Ile	Asn	Lys	Gly	Gly	Gln	Leu	Met	Asn	Thr	Asn	
		35					40					45				
Glu	Ile	Glu	Asn	Trp	Pro	Gly	Asp	Phe	Lys	Lys	Ile	Thr	Gly	Pro	Glu	
	50					55					60					
Leu	Met	Asn	Arg	Met	His	Glu	His	Ser	Leu	Lys	Phe	Lys	Thr	Glu	Ile	
65					70					75					80	
Val	Tyr	Asp	Asn	Ile	Ile	Ser	Val	Glu	Phe	Lys	Lys	Lys	Pro	Phe	Phe	
			85						90					95		
Leu	Leu	Gly	Glu	Tyr	Asn	Lys	Tyr	Thr	Cys	Asp	Ala	Val	Ile	Ile	Ala	
			100					105					110			
Thr	Gly	Ala	Asn	Pro	Arg	Tyr	Leu	Gly	Leu	Ser	Ser	Glu	Asn	Lys	Phe	
		115					120					125				
Lys	Gly	Lys	Gly	Ile	Ser	Thr	Cys	Ala	Val	Cys	Asp	Gly	Phe	Phe	Tyr	
	130					135					140					
Lys	Asn	Lys	Glu	Ile	Ala	Val	Val	Gly	Gly	Gly	Asn	Thr	Ala	Ile	Glu	
145					150					155					160	
Glu	Thr	Leu	Tyr	Leu	Ser	Asn	Phe	Val	Lys	Lys	Ile	Tyr	Leu	Ile	His	
				165					170					175		
Arg	Arg	Asn	Asn	Phe	Lys	Ala	Glu	Lys	Ile	Leu	Ile	Asp	Arg	Leu	Leu	
		180						185					190			

Lys Ile Val Lys Thr Lys Lys Val Ile Leu His Leu Asn Ser Thr Ile
 195 200 205
 Glu Asp Ile Leu Gly Asn Asn Lys Gly Val Thr His Leu Leu Ile Lys
 210 215 220
 Asn Lys Asn Leu Lys Glu Lys Lys Lys Leu Lys Ile Ala Val Ser Gly
 225 230 235 240
 Leu Phe Val Ala Ile Gly Tyr Ile Pro Asn Thr Asp Ile Phe Thr Asp
 245 250 255
 Gln Leu Lys Met Lys Asp Gly Tyr Ile Lys Ile Lys Lys Gly Thr His
 260 265 270
 Gly Asn Tyr Thr Gln Thr Asn Ile Pro Gly Val Phe Ala Ala Gly Asp
 275 280 285
 Val Ile Asp His Val Tyr Arg Gln Ala Ile Thr Ser Ser Ala Ser Gly
 290 295 300
 Cys Met Ala Ala Leu Asp Ser Glu Arg Tyr Leu Asn Ser Leu Ser
 305 310 315

<210> 202

<211> 312

<212> PRT

<213> Chlamydia muridarum

<400> 202

Met Thr His Val Lys Leu Ala Ile Ile Gly Ser Gly Pro Ala Gly Tyr
 1 5 10 15
 Thr Ala Ala Ile Tyr Ala Ser Arg Ala Leu Leu Thr Pro Ile Leu Phe
 20 25 30
 Glu Gly Phe Phe Ser Gly Ile Ala Gly Gly Gln Leu Met Thr Thr Thr
 35 40 45
 Glu Val Glu Asn Phe Pro Gly Phe Pro Gln Gly Val Leu Gly His Gln
 50 55 60
 Leu Met Glu Asn Met Lys Met Gln Ala Gln Arg Phe Gly Thr Gln Val
 65 70 75 80
 Ile Ala Lys Asp Ile Thr Ser Val Asp Phe Ser Val Arg Pro Phe Val
 85 90 95
 Leu Lys Ser Gly Glu Asp Thr Phe Thr Cys Asp Ala Cys Ile Ile Ala
 100 105 110
 Thr Gly Ala Ser Ala Lys Arg Leu Ser Ile Pro Gly Ala Gly Asp Asn
 115 120 125
 Glu Phe Trp Gln Lys Gly Val Thr Ala Cys Ala Val Cys Asp Gly Ala
 130 135 140
 Ser Pro Ile Phe Arg Asp Arg Asp Leu Phe Val Ile Gly Gly Gly Asp
 145 150 155 160
 Ser Ala Leu Glu Glu Ala Met Phe Leu Thr Arg Tyr Gly Lys Arg Val
 165 170 175
 Phe Val Val His Arg Arg Asp Thr Leu Arg Ala Ser Lys Ala Met Val
 180 185 190
 Asn Lys Ala Gln Ala Asn Glu Lys Ile Val Phe Leu Trp Asn Ser Glu
 195 200 205
 Val Val Lys Ile Leu Gly Asp Ser Leu Val Arg Ser Ile Asp Ile Phe
 210 215 220
 Asn Asn Val Glu Lys Thr Thr Val Thr Met Glu Ala Ala Gly Val Phe
 225 230 235 240
 Phe Ala Ile Gly His Gln Pro Asn Thr Ala Phe Leu Gly Gly Gln Leu
 245 250 255
 Ser Leu Asp Glu Asn Gly Tyr Ile Ile Thr Glu Lys Gly Ser Ser Arg
 260 265 270
 Thr Ser Val Pro Gly Val Phe Ala Ala Gly Asp Val Gln Asp Lys Tyr
 275 280 285
 Tyr Arg Gln Ala Ile Thr Ser Ala Gly Ser Gly Cys Met Ala Ala Leu
 290 295 300
 Asp Ala Glu Arg Phe Leu Glu Lys
 305 310

<210> 203

<211> 311

<212> PRT

<213> Chlamydia pneumoniae

<400> 203

```

Met Ile His Ser Arg Leu Ile Ile Ile Gly Ser Gly Pro Ser Gly Tyr
 1           5           10           15
Thr Ala Ala Ile Tyr Ala Ser Arg Ala Leu Leu His Pro Leu Leu Phe
           20           25           30
Glu Gly Phe Phe Ser Gly Ile Ser Gly Gly Gln Leu Met Thr Thr Thr
           35           40           45
Glu Val Glu Asn Phe Pro Gly Phe Pro Glu Gly Ile Leu Gly Pro Lys
 50           55           60
Leu Met Asn Asn Met Lys Glu Gln Ala Val Arg Phe Gly Thr Lys Thr
65           70           75           80
Leu Ala Gln Asp Ile Ile Ser Val Asp Phe Ser Val Arg Pro Phe Ile
           85           90           95
Leu Lys Ser Lys Glu Glu Thr Tyr Ser Cys Asp Ala Cys Ile Ile Ala
           100          105          110
Thr Gly Ala Ser Ala Lys Arg Leu Glu Ile Pro Gly Ala Gly Asn Asp
           115          120          125
Glu Phe Trp Gln Lys Gly Val Thr Ala Cys Ala Val Cys Asp Gly Ala
130          135          140
Ser Pro Ile Phe Lys Asn Lys Asp Leu Tyr Val Ile Gly Gly Gly Asp
145          150          155          160
Ser Ala Leu Glu Glu Ala Leu Tyr Leu Thr Arg Tyr Gly Ser His Val
           165          170          175
Tyr Val Val His Arg Arg Asp Lys Leu Arg Ala Ser Lys Ala Met Glu
           180          185          190
Ala Arg Ala Gln Asn Asn Glu Lys Ile Thr Phe Leu Trp Asn Ser Glu
           195          200          205
Ile Val Lys Ile Ser Gly Asp Ser Ile Val Arg Ser Val Asp Ile Lys
210          215          220
Asn Val Gln Thr Gln Glu Ile Thr Thr Arg Glu Ala Ala Gly Val Phe
225          230          235          240
Phe Ala Ile Gly His Lys Pro Asn Thr Asp Phe Leu Gly Gly Gln Leu
           245          250          255
Thr Leu Asp Glu Ser Gly Tyr Ile Val Thr Glu Lys Gly Thr Ser Lys
           260          265          270
Thr Ser Val Pro Gly Val Phe Ala Ala Gly Asp Val Gln Asp Lys Tyr
           275          280          285
Tyr Arg Gln Ala Val Thr Ser Ala Gly Ser Gly Cys Ile Ala Ala Leu
290          295          300
Asp Ala Glu Arg Phe Leu Gly
305          310

```

<210> 204

<211> 312

<212> PRT

<213> Chlamydia trachomatis

<400> 204

```

Met Thr His Ala Lys Leu Val Ile Ile Gly Ser Gly Pro Ala Gly Tyr
 1           5           10           15
Thr Ala Ala Ile Tyr Ala Ser Arg Ala Leu Leu Thr Pro Val Leu Phe
           20           25           30
Glu Gly Phe Phe Ser Gly Ile Ala Gly Gly Gln Leu Met Thr Thr Thr
           35           40           45
Glu Val Glu Asn Phe Pro Gly Phe Pro Glu Gly Val Leu Gly His Gln
 50           55           60
Leu Met Asp Leu Met Lys Thr Gln Ala Gln Arg Phe Gly Thr Gln Val
65           70           75           80
Leu Ser Lys Asp Ile Thr Ala Val Asp Phe Ser Val Arg Pro Phe Val
           85           90           95
Leu Lys Ser Gly Lys Glu Thr Phe Thr Cys Asp Ala Cys Ile Ile Ala
           100          105          110

```

Thr Gly Ala Ser Ala Lys Arg Leu Ser Ile Pro Gly Ala Gly Asp Asn
 115 120 125
 Glu Phe Trp Gln Lys Gly Val Thr Ala Cys Ala Val Cys Asp Gly Ala
 130 135 140
 Ser Pro Ile Phe Arg Asp Lys Asp Leu Phe Val Val Gly Gly Gly Asp
 145 150 155 160
 Ser Ala Leu Glu Glu Ala Met Phe Leu Thr Arg Tyr Gly Lys Arg Val
 165 170 175
 Phe Val Val His Arg Arg Asp Thr Leu Arg Ala Ser Lys Val Met Val
 180 185 190
 Asn Lys Ala Gln Ala Asn Glu Lys Ile Phe Phe Leu Trp Asn Ser Glu
 195 200 205
 Ile Val Lys Ile Ser Gly Asp Thr Leu Val Arg Ser Ile Asp Ile Tyr
 210 215 220
 Asn Asn Val Asp Glu Thr Thr Thr Thr Met Glu Ala Ala Gly Val Phe
 225 230 235 240
 Phe Ala Ile Gly His Gln Pro Asn Thr Ala Phe Leu Gly Gly Gln Val
 245 250 255
 Ala Leu Asp Glu Asn Gly Tyr Ile Ile Thr Glu Lys Gly Ser Ser Arg
 260 265 270
 Thr Ser Val Pro Gly Val Phe Ala Ala Gly Asp Val Gln Asp Lys Tyr
 275 280 285
 Tyr Arg Gln Ala Ile Thr Ser Ala Gly Ser Gly Cys Met Ala Ala Leu
 290 295 300
 Asp Ala Glu Arg Phe Leu Glu Asn
 305 310

<210> 205

<211> 315

<212> PRT

<213> Clostridium litorale

<400> 205

Met Glu Asn Val Tyr Asp Ile Ala Ile Ile Gly Ser Gly Pro Ala Gly
 1 5 10 15
 Leu Ala Ala Ala Leu Tyr Gly Ala Arg Ala Lys Met Lys Thr Leu Leu
 20 25 30
 Leu Glu Gly Met Lys Val Gly Gly Gln Ile Val Ile Thr His Glu Val
 35 40 45
 Ala Asn Tyr Pro Gly Ser Val Pro Glu Ala Thr Gly Pro Ser Leu Ile
 50 55 60
 Gly Arg Met Glu Glu Gln Val Glu Glu Phe Gly Ala Glu Arg Val Met
 65 70 75 80
 Asp Asn Ile Val Asp Val Asp Phe Thr Asp Lys Ile Lys Val Leu Lys
 85 90 95
 Gly Ala Lys Gly Glu Tyr Lys Ala Lys Ala Val Ile Val Ala Thr Gly
 100 105 110
 Ala Ser Pro Lys Leu Ala Gly Cys Pro Gly Glu Lys Glu Leu Thr Gly
 115 120 125
 Lys Gly Val Ser Tyr Cys Ala Thr Cys Asp Ala Asp Phe Phe Glu Asp
 130 135 140
 Met Glu Val Phe Val Ile Gly Gly Gly Asp Thr Ala Val Glu Glu Ala
 145 150 155 160
 Met Phe Leu Thr Lys Phe Ala Arg Lys Val Thr Ile Val His Arg Arg
 165 170 175
 Ala Glu Leu Arg Ala Ala Lys Ser Ile Gln Glu Lys Ala Phe Lys Asn
 180 185 190
 Glu Lys Leu Asn Phe Met Trp Asn Thr Val Ile Glu Glu Ile Lys Gly
 195 200 205
 Asp Gly Ile Val Glu Ser Ala Val Phe Lys Asn Arg Glu Thr Gly Glu
 210 215 220
 Val Thr Glu Phe Val Ala Pro Glu Glu Asp Gly Thr Phe Gly Ile Phe
 225 230 235 240
 Val Phe Ile Gly Tyr Asp Pro Lys Ser Ala Leu Val Glu Gly Lys Leu
 245 250 255
 Glu Leu Asp Glu Thr Gly Tyr Ile Pro Thr Asp Asp Asn Met Lys Thr

260 265 270
 Asn Val Glu Gly Val Phe Ala Ala Gly Asp Ile Arg Val Lys Ser Leu
 275 280 285
 Arg Gln Val Val Thr Ala Thr Ala Asp Gly Ala Ile Ala Ala Val Gln
 290 295 300
 Ala Glu Lys Tyr Ile Glu Glu Leu Phe Ala Glu
 305 310 315

<210> 206
 <211> 321
 <212> PRT
 <213> *Coxiella burnetii*

<400> 206
 Met Asn Lys Pro Gln His His Ser Leu Ile Ile Leu Gly Ser Gly Pro
 1 5 10 15
 Ala Gly Tyr Thr Asp Ala Ile Tyr Val Ala Arg Ala Asn Leu Lys Pro
 20 25 30
 Ile Met Ile Thr Gly Met Glu Gln Gly Gly Gln Leu Met Thr Thr Thr
 35 40 45
 Asp Val Ala Asn Trp Pro Gly Glu Ala Pro Gly Leu Gln Gly Pro Lys
 50 55 60
 Leu Leu Glu Arg Met Gln Lys His Ala Gly Gly Ala Leu Asn Thr Gln
 65 70 75 80
 Phe Ile Phe Asp His Ile Asn Lys Pro Asp Leu Asn Pro Arg Pro Phe
 85 90 95
 Leu Leu Gln Gly Asp Asn Ala Thr Tyr Ser Cys Asp Ala Leu Ile Ile
 100 105 110
 Ala Thr Gly Ala Ser Ala Arg Tyr Leu Gly Leu Pro Ser Glu Lys Pro
 115 120 125
 Tyr Met Gly Lys Gly Val Ser Ala Cys Ala Thr Cys Asp Gly Phe Phe
 130 135 140
 Tyr Arg Ala Lys Lys Val Ala Val Val Gly Gly Gly Asn Thr Ser Val
 145 150 155 160
 Glu Glu Ala Leu Tyr Leu Ser His Ile Ala Ser His Val Thr Leu Ile
 165 170 175
 His Arg Arg Asp Lys Leu Arg Ala Glu Lys Met Leu Ser Ala Gln Leu
 180 185 190
 Ile Lys Lys Val Glu Glu Gly Lys Val Ala Ile Val Trp Ser His Val
 195 200 205
 Ile Glu Glu Val Leu Gly Asp Asp Gln Gly Val Thr Gly Val His Leu
 210 215 220
 Lys His Val Lys Glu Glu Lys Thr Gln Asp Leu Thr Ile Asp Gly Leu
 225 230 235 240
 Phe Ile Ala Ile Gly His Asp Pro Asn Thr Lys Ile Phe Lys Glu Gln
 245 250 255
 Leu Glu Met Asp Glu Ala Gly Tyr Leu Arg Ala Lys Ser Gly Leu Gln
 260 265 270
 Gly Asn Ala Thr Ala Thr Asn Ile Pro Gly Val Phe Pro Ala Val Val
 275 280 285
 Val Arg Gly Gln Leu Tyr Arg Gln Thr Ile Ala Ala Gly Met Gly
 290 295 300
 Cys Met Pro Ala Leu Asp Ala Glu Arg Tyr Leu Asp Ser Leu Asn Gln
 305 310 315 320
 Ala

<210> 207
 <211> 320
 <212> PRT
 <213> *Escherichia coli*

<400> 207
 Gly Thr Thr Lys His Ser Lys Leu Leu Ile Leu Gly Ser Gly Pro Ala
 1 5 10 15

Gly Tyr Thr Ala Ala Val Tyr Ala Ala Arg Ala Asn Leu Gln Pro Val
 20 25 30
 Leu Ile Thr Gly Met Glu Lys Gly Gly Gln Leu Thr Thr Thr Glu
 35 40 45
 Val Glu Asn Trp Pro Gly Asp Pro Asn Asp Leu Thr Gly Pro Leu Leu
 50 55 60
 Met Glu Arg Met His Glu His Ala Thr Lys Phe Glu Thr Glu Ile Ile
 65 70 75 80
 Phe Asp His Ile Asn Lys Val Asp Leu Gln Asn Arg Pro Phe Arg Leu
 85 90 95
 Asn Gly Asp Asn Gly Glu Tyr Thr Cys Asp Ala Leu Ile Ile Ala Thr
 100 105 110
 Gly Ala Ser Ala Arg Tyr Leu Gly Leu Pro Ser Glu Glu Ala Phe Lys
 115 120 125
 Gly Arg Gly Val Ser Ala Cys Ala Thr Cys Asp Gly Phe Phe Tyr Arg
 130 135 140
 Asn Gln Lys Val Ala Val Ile Gly Gly Gly Asn Thr Ala Val Glu Glu
 145 150 155 160
 Ala Leu Tyr Leu Ser Asn Ile Ala Ser Glu Val His Leu Ile His Arg
 165 170 175
 Arg Asp Gly Phe Arg Ala Glu Lys Ile Leu Ile Lys Arg Leu Met Asp
 180 185 190
 Lys Val Glu Asn Gly Asn Ile Ile Leu His Thr Asn Arg Thr Leu Glu
 195 200 205
 Glu Val Thr Gly Asp Gln Met Gly Val Thr Gly Val Arg Leu Arg Asp
 210 215 220
 Thr Gln Asn Ser Asp Asn Ile Glu Ser Leu Asp Val Ala Gly Leu Phe
 225 230 235 240
 Val Ala Ile Gly His Ser Pro Asn Thr Ala Ile Phe Glu Gly Gln Leu
 245 250 255
 Glu Leu Glu Asn Gly Tyr Ile Lys Val Gln Ser Gly Ile His Gly Asn
 260 265 270
 Ala Thr Gln Thr Ser Ile Pro Gly Val Phe Ala Ala Gly Asp Val Met
 275 280 285
 Asp His Ile Tyr Arg Gln Ala Ile Thr Ser Ala Gly Thr Gly Cys Met
 290 295 300
 Ala Ala Leu Asp Ala Glu Arg Tyr Leu Asp Gly Leu Ala Asp Ala Lys
 305 310 315 320

<210> 208

<211> 315

<212> PRT

<213> Eubacterium acidaminophilum

<400> 208

Met Glu Asn Val Tyr Asp Leu Ala Ile Ile Gly Ser Gly Pro Ala Gly
 1 5 10 15
 Leu Ala Ala Ala Leu Tyr Gly Ala Arg Ala Lys Met Lys Thr Ile Met
 20 25 30
 Ile Glu Gly Gln Lys Val Gly Gly Gln Ile Val Ile Thr His Glu Val
 35 40 45
 Ala Asn Tyr Pro Gly Ser Val Arg Glu Ala Thr Gly Pro Ser Leu Ile
 50 55 60
 Glu Arg Met Glu Glu Gln Ala Asn Glu Phe Gly Ala Glu Lys Val Met
 65 70 75 80
 Asp Lys Ile Val Asp Val Asp Leu Asp Gly Lys Ile Lys Val Ile Lys
 85 90 95
 Gly Glu Lys Ala Glu Tyr Lys Ala Lys Ser Val Ile Leu Ala Thr Gly
 100 105 110
 Ala Ala Pro Arg Leu Ala Gly Cys Pro Gly Glu Gln Glu Leu Thr Gly
 115 120 125
 Lys Gly Val Ser Tyr Cys Ala Thr Cys Asp Ala Asp Phe Phe Glu Asp
 130 135 140
 Met Glu Val Phe Val Val Gly Gly Gly Asp Thr Ala Val Glu Glu Ala
 145 150 155 160
 Met Tyr Leu Ala Lys Phe Ala Arg Lys Val Thr Ile Val His Arg Arg

Asp	Glu	Leu	Arg	165	Ala	Ala	Lys	Ser	Ile	Gln	Glu	Lys	Ala	Phe	Lys	Asn	175
Pro	Lys	Leu	Asp	180	Phe	Met	Trp	Asn	Ser	Ala	Ile	Glu	Glu	Ile	Lys	Gly	190
Asp	Gly	Ile	Val	195	Glu	Ser	Ala	Val	Phe	Lys	Asn	Leu	Val	Thr	Gly	Glu	205
Thr	Thr	Glu	Tyr	210	Phe	Ala	Asn	Glu	Glu	Asp	Gly	Thr	Phe	Gly	Ile	Phe	220
225	Val	Phe	Ile	230	Gly	Tyr	Ile	Pro	Lys	Ser	Asp	Val	Phe	Lys	Gly	Lys	235
Val	Phe	Ile	Gly	245	Tyr	Ile	Pro	Lys	Ser	Asp	Val	Phe	Lys	Gly	Lys	Ile	250
Thr	Leu	Asp	Asp	260	Ala	Gly	Tyr	Ile	Ile	Thr	Asp	Asp	Asn	Met	Lys	Thr	265
Asn	Val	Glu	Gly	275	Val	Phe	Ala	Ala	Gly	Asp	Ile	Arg	Val	Lys	Ser	Leu	280
Arg	Gln	Val	Val	290	Thr	Ala	Cys	Ala	Asp	Gly	Ala	Ile	Ala	Ala	Thr	Gln	295
Ala	Glu	Lys	Tyr	305	Val	Glu	Ala	Asn	Phe	Glu	Glu						310

<210> 209

<211> 318

<212> PRT

<213> Haemophilus influenzae

<400> 209

Met	Ser	Asp	Ile	Lys	His	Ala	Lys	Leu	Leu	Ile	Leu	Gly	Ser	Gly	Pro	
1				5				10						15		
Ala	Gly	Tyr	Thr	Ala	Ala	Ile	Tyr	Ala	Ala	Arg	Ala	Asn	Leu	Lys	Pro	
			20					25					30			
Val	Leu	Val	Thr	Gly	Leu	Gln	Gln	Gly	Gly	Gln	Leu	Thr	Thr	Thr	Asp	
			35				40					45				
Glu	Ile	Glu	Asn	Trp	Pro	Gly	Asp	Phe	Glu	Met	Thr	Thr	Gly	Ser	Gly	
50						55					60					
Leu	Met	Gln	Arg	Met	Leu	Gln	His	Ala	Glu	Lys	Phe	Glu	Thr	Glu	Ile	
65					70					75					80	
Val	Phe	Asp	His	Ile	Asn	Arg	Val	Asp	Leu	Ser	Ser	Arg	Pro	Phe	Lys	
			85						90					95		
Leu	Phe	Gly	Asp	Val	Gln	Asn	Phe	Thr	Cys	Asp	Ala	Leu	Ile	Ile	Ala	
			100					105					110			
Thr	Gly	Ala	Ser	Ala	Arg	Tyr	Ile	Gly	Leu	Pro	Ser	Glu	Glu	Asn	Tyr	
			115				120					125				
Lys	Gly	Arg	Gly	Val	Ser	Ala	Cys	Ala	Thr	Cys	Asp	Gly	Phe	Phe	Tyr	
			130			135					140					
Arg	Asn	Lys	Pro	Val	Gly	Val	Ile	Gly	Gly	Gly	Asn	Thr	Ala	Val	Glu	
145					150					155					160	
Glu	Ala	Leu	Tyr	Leu	Ala	Asn	Ile	Ala	Ser	Thr	Val	His	Leu	Ile	His	
				165					170					175		
Arg	Arg	Asp	Ser	Phe	Arg	Ala	Glu	Lys	Ile	Leu	Ile	Asp	Arg	Leu	Tyr	
			180					185					190			
Lys	Lys	Val	Glu	Glu	Gly	Lys	Ile	Val	Leu	His	Thr	Asp	Arg	Thr	Leu	
			195				200					205				
Asp	Glu	Val	Leu	Gly	Asp	Asn	Met	Gly	Val	Thr	Gly	Leu	Arg	Leu	Ala	
			210			215					220					
Asn	Thr	Lys	Thr	Gly	Glu	Lys	Glu	Glu	Leu	Lys	Leu	Asp	Gly	Leu	Phe	
225					230					235					240	
Val	Ala	Ile	Gly	His	Ser	Pro	Asn	Thr	Glu	Ile	Phe	Gln	Gly	Gln	Leu	
				245					250					255		
Glu	Leu	Asn	Asn	Gly	Tyr	Ile	Val	Val	Lys	Ser	Gly	Leu	Asp	Gly	Asn	
			260					265					270			
Ala	Thr	Ala	Thr	Ser	Val	Glu	Gly	Val	Phe	Ala	Ala	Gly	Asp	Val	Met	
			275				280					285				
Asp	His	Asn	Tyr	Arg	Gln	Ala	Ile	Thr	Ser	Ala	Gly	Thr	Gly	Cys	Met	
	290					295					300					
Ala	Ala	Leu	Asp	Ala	Glu	Arg	Tyr	Leu	Asp	Ala	Gln	Glu	Ala			
305					310						315					

<210> 210
 <211> 311
 <212> PRT
 <213> Helicobacter pylori

<400> 210
 Met Ile Asp Cys Ala Ile Ile Gly Gly Gly Pro Ala Gly Leu Ser Ala
 1 5 10 15
 Gly Leu Tyr Ala Thr Arg Gly Gly Val Lys Asn Ala Val Leu Phe Glu
 20 25 30
 Lys Gly Met Pro Gly Gly Gln Ile Thr Gly Ser Ser Glu Ile Glu Asn
 35 40 45
 Tyr Pro Gly Val Lys Glu Val Val Ser Gly Leu Asp Phe Met Gln Pro
 50 55 60
 Trp Gln Glu Gln Cys Phe Arg Phe Gly Leu Lys His Glu Met Thr Ala
 65 70 75 80
 Ile Gln Arg Val Ser Lys Lys Gly Ser His Phe Val Ile Leu Ala Glu
 85 90 95
 Asp Gly Lys Thr Phe Glu Ala Lys Ser Val Ile Ile Ala Thr Gly Gly
 100 105 110
 Ser Pro Lys Arg Thr Gly Ile Lys Gly Glu Ser Glu Tyr Trp Gly Lys
 115 120 125
 Gly Val Ser Thr Cys Ala Thr Cys Asp Gly Phe Phe Tyr Lys Asn Lys
 130 135 140
 Glu Val Ala Val Leu Gly Gly Gly Asp Thr Ala Val Glu Glu Ala Ile
 145 150 155 160
 Tyr Leu Ala Asn Ile Cys Lys Lys Val Tyr Leu Ile His Arg Arg Asp
 165 170 175
 Gly Phe Arg Cys Ala Pro Ile Thr Leu Glu His Ala Lys Asn Asn Ser
 180 185 190
 Lys Ile Glu Phe Leu Thr Pro Tyr Val Val Glu Glu Ile Lys Gly Asp
 195 200 205
 Ala Ser Gly Val Ser Ser Leu Ser Ile Lys Asn Thr Ala Thr Asn Glu
 210 215 220
 Lys Arg Glu Leu Val Val Pro Gly Leu Phe Ile Phe Val Gly Tyr Asp
 225 230 235 240
 Val Asn Asn Ala Val Leu Lys Gln Glu Asp Asn Ser Met Leu Cys Glu
 245 250 255
 Cys Asp Glu Tyr Gly Ser Ile Val Val Asp Phe Ser Met Lys Thr Asn
 260 265 270
 Val Gln Gly Leu Phe Ala Ala Gly Asp Ile Arg Ile Phe Ala Pro Lys
 275 280 285
 Gln Val Val Cys Ala Ala Ser Asp Gly Ala Thr Ala Ala Leu Ser Val
 290 295 300
 Ile Ser Tyr Leu Glu His His
 305 310

<210> 211
 <211> 311
 <212> PRT
 <213> Helicobacter pylori

<400> 211
 Met Ile Asp Cys Ala Ile Ile Gly Gly Gly Pro Ala Gly Leu Ser Ala
 1 5 10 15
 Gly Leu Tyr Ala Thr Arg Gly Gly Val Lys Asn Ala Val Leu Phe Glu
 20 25 30
 Lys Gly Met Pro Gly Gly Gln Ile Thr Gly Ser Ser Glu Ile Glu Asn
 35 40 45
 Tyr Pro Gly Val Lys Glu Val Val Ser Gly Leu Asp Phe Met Gln Pro
 50 55 60
 Trp Gln Glu Gln Cys Phe Arg Phe Gly Leu Lys His Glu Met Thr Ala
 65 70 75 80
 Val Gln Arg Val Ser Lys Lys Asp Ser His Phe Val Ile Leu Ala Glu
 85 90 95

```

Asp Gly Lys Thr Phe Glu Ala Lys Ser Val Ile Ile Ala Thr Gly Gly
    100      105      110
Ser Pro Lys Arg Thr Gly Ile Lys Gly Glu Ser Glu Tyr Trp Gly Lys
    115      120      125
Gly Val Ser Thr Cys Ala Thr Cys Asp Gly Phe Phe Tyr Lys Asn Lys
    130      135      140
Glu Val Ala Val Leu Gly Gly Asp Thr Ala Val Glu Glu Ala Ile
    145      150      155      160
Tyr Leu Ala Asn Ile Cys Lys Lys Val Tyr Leu Ile His Arg Arg Asp
    165      170      175
Gly Phe Arg Cys Ala Pro Ile Thr Leu Glu His Ala Lys Asn Asn Asp
    180      185      190
Lys Ile Glu Phe Leu Thr Pro Tyr Val Val Glu Glu Ile Lys Gly Asp
    195      200      205
Ala Ser Gly Val Ser Ser Leu Ser Ile Lys Asn Thr Ala Thr Asn Glu
    210      215      220
Lys Arg Glu Leu Val Val Pro Gly Phe Phe Ile Phe Val Gly Tyr Asp
    225      230      235      240
Val Asn Asn Ala Val Leu Lys Gln Glu Asp Asn Ser Met Leu Cys Lys
    245      250      255
Cys Asp Glu Tyr Gly Ser Ile Val Val Asp Phe Ser Met Lys Thr Asn
    260      265      270
Val Gln Gly Leu Phe Ala Ala Gly Asp Ile Arg Ile Phe Ala Pro Lys
    275      280      285
Gln Val Val Cys Ala Ala Ser Asp Gly Ala Thr Ala Ala Leu Ser Val
    290      295      300
Ile Ser Tyr Leu Glu His His
    305      310

```

<210> 212

<211> 319

<212> PRT

<213> *Listeria monocytogenes*

<400> 212

```

Met Ala Ser Glu Glu Lys Ile Tyr Asp Val Ile Ile Ile Gly Ala Gly
  1      5      10      15
Pro Ala Gly Met Thr Ala Ala Leu Tyr Thr Ser Arg Ala Asp Leu Asp
    20      25      30
Thr Leu Met Ile Glu Arg Gly Val Pro Gly Gly Gln Met Val Asn Thr
    35      40      45
Ala Glu Val Glu Asn Tyr Pro Gly Phe Asp Ser Ile Leu Gly Pro Asp
    50      55      60
Leu Ser Asp Lys Met Leu Ser Gly Ala Lys Gln Phe Gly Ala Glu Tyr
    65      70      75      80
Ala Tyr Gly Asp Ile Lys Glu Val Val Asp Gly Lys Glu Phe Lys Thr
    85      90      95
Val Thr Ala Gly Ser Lys Thr Tyr Lys Ala Arg Ala Ile Ile Ile Ala
    100      105      110
Thr Gly Ala Glu His Arg Lys Leu Gly Ala Ala Gly Glu Glu Leu
    115      120      125
Ser Gly Arg Gly Val Ser Tyr Cys Ala Val Cys Asp Gly Ala Phe Phe
    130      135      140
Lys Asn Arg Glu Leu Ile Val Val Gly Gly Gly Asp Ser Ala Val Glu
    145      150      155      160
Glu Gly Thr Tyr Leu Thr Arg Tyr Ala Asp Lys Val Thr Ile Val His
    165      170      175
Arg Arg Asp Lys Leu Arg Ala Gln Gln Ile Leu Gln Asp Arg Ala Phe
    180      185      190
Lys Asp Glu Lys Val Asp Phe Ile Trp Asn Ser Thr Val Glu Glu Ile
    195      200      205
Val Gly Asp Gly Lys Lys Val Thr Gly Ala Lys Leu Val Ser Thr Val
    210      215      220
Asp Gly Ser Glu Ser Ile Met Pro Val Asp Gly Val Phe Ile Tyr Val
    225      230      235      240
Gly Leu Val Pro Leu Thr Lys Ala Phe Leu Asn Leu Gly Ile Thr Asp

```

Asp	Glu	Gly	Tyr	Ile	Val	Thr	Asp	Glu	Glu	Met	Arg	Thr	Asn	Leu	Pro
			245						250					255	
Gly	Ile	Phe	Ala	Ala	Gly	Asp	Val	Arg	Ala	Lys	Ser	Leu	Arg	Gln	Ile
		260							265					270	
Val	Thr	Ala	Thr	Gly	Asp	Gly	Gly	Leu	Ala	Gly	Gln	Asn	Ala	Gln	Lys
	275						280					285			
Tyr	Val	Glu	Glu	Leu	Lys	Glu	Ser	Leu	Glu	Ala	Glu	Ala	Ala	Lys	
	290					295					300				
305					310					315					

<210> 213

<211> 315

<212> PRT

<213> Mycoplasma genitalium

<400> 213

Met	Leu	Lys	Val	Asn	Ala	Asp	Phe	Leu	Thr	Lys	Asp	Gln	Val	Ile	Tyr
1				5					10					15	
Asp	Leu	Val	Ile	Val	Gly	Ala	Gly	Pro	Ala	Gly	Ile	Ala	Ser	Ala	Ile
		20						25					30		
Tyr	Gly	Lys	Arg	Ala	Asn	Leu	Asn	Leu	Ala	Ile	Ile	Glu	Gly	Asn	Thr
	35						40					45			
Pro	Gly	Gly	Lys	Ile	Val	Lys	Thr	Asn	Ile	Val	Glu	Asn	Tyr	Pro	Gly
	50					55					60				
Phe	Lys	Thr	Ile	Thr	Gly	Pro	Glu	Leu	Gly	Leu	Glu	Met	Tyr	Asn	His
65					70					75					80
Leu	Leu	Ala	Phe	Glu	Pro	Val	Val	Phe	Tyr	Asn	Asn	Leu	Ile	Lys	Ile
			85					90						95	
Asp	His	Leu	Asn	Asp	Thr	Phe	Ile	Leu	Tyr	Leu	Asp	Asn	Lys	Thr	Thr
			100					105					110		
Val	Phe	Ser	Lys	Thr	Val	Ile	Tyr	Ala	Thr	Gly	Met	Glu	Glu	Arg	Lys
	115						120					125			
Leu	Gly	Ile	Glu	Lys	Glu	Asp	Tyr	Phe	Tyr	Gly	Lys	Gly	Ile	Ser	Tyr
	130					135					140				
Cys	Ala	Ile	Cys	Asp	Ala	Ala	Leu	Tyr	Lys	Gly	Lys	Thr	Val	Gly	Val
145					150					155					160
Val	Gly	Gly	Gly	Asn	Ser	Ala	Ile	Gln	Glu	Ala	Ile	Tyr	Leu	Ser	Ser
				165					170					175	
Ile	Ala	Lys	Thr	Val	His	Leu	Ile	His	Arg	Arg	Glu	Val	Phe	Arg	Ser
			180					185					190		
Asp	Ala	Leu	Leu	Val	Glu	Lys	Leu	Lys	Lys	Ile	Ser	Asn	Val	Val	Phe
	195						200					205			
His	Leu	Asn	Ala	Thr	Val	Lys	Gln	Leu	Ile	Gly	Gln	Glu	Lys	Leu	Gln
	210					215					220				
Thr	Val	Lys	Leu	Ala	Ser	Thr	Val	Asp	Lys	Ser	Glu	Ser	Glu	Ile	Ala
225					230					235					240
Ile	Asp	Cys	Leu	Phe	Pro	Tyr	Ile	Gly	Phe	Glu	Ser	Asn	Asn	Lys	Pro
				245					250					255	
Val	Leu	Asp	Leu	Lys	Leu	Asn	Leu	Asp	Gln	Asn	Gly	Phe	Ile	Leu	Gly
	260							265					270		
Asp	Glu	Asn	Met	Gln	Thr	Asn	Ile	Lys	Gly	Phe	Tyr	Val	Ala	Gly	Asp
	275						280					285			
Cys	Arg	Ser	Lys	Ser	Phe	Arg	Gln	Ile	Ala	Thr	Ala	Ile	Ser	Asp	Gly
	290					295					300				
Val	Thr	Ala	Val	Leu	Lys	Val	Arg	Asp	Asp	Ile					
305					310					315					

<210> 214

<211> 458

<212> PRT

<213> Mycobacterium leprae

<400> 214

Met	Asn	Thr	Thr	Pro	Ser	Ala	His	Glu	Thr	Ile	His	Glu	Val	Ile	Val
1				5					10					15	

```

Ile Gly Ser Gly Pro Ala Gly Tyr Thr Ala Ala Leu Tyr Ala Ala Arg
      20      25      30
Ala Gln Leu Thr Pro Leu Val Phe Glu Gly Thr Ser Phe Gly Gly Ala
      35      40      45
Leu Met Thr Thr Thr Glu Val Glu Asn Tyr Pro Gly Phe Arg Asn Gly
      50      55      60
Ile Thr Gly Pro Glu Leu Met Asp Asp Met Arg Glu Gln Ala Leu Arg
      65      70      75
Phe Gly Ala Glu Leu Arg Thr Glu Asp Val Glu Ser Val Ser Leu Arg
      85      90      95
Gly Pro Ile Lys Ser Val Val Thr Ala Glu Gly Gln Thr Tyr Gln Ala
      100     105     110
Arg Ala Val Ile Leu Ala Met Gly Thr Ser Val Arg Tyr Leu Gln Ile
      115     120     125
Pro Gly Glu Gln Glu Leu Leu Gly Arg Gly Val Ser Ala Cys Ala Thr
      130     135     140
Cys Asp Gly Ser Phe Phe Arg Gly Gln Asp Ile Ala Val Ile Gly Gly
      145     150     155
Gly Asp Ser Ala Met Glu Glu Ala Leu Phe Leu Thr Arg Phe Ala Arg
      165     170     175
Ser Val Thr Leu Val His Arg Arg Asp Glu Phe Arg Ala Ser Lys Ile
      180     185     190
Met Leu Gly Arg Ala Arg Asn Asn Asp Lys Ile Lys Phe Ile Thr Asn
      195     200     205
His Thr Val Val Ala Val Asn Gly Tyr Thr Thr Val Thr Gly Leu Arg
      210     215     220
Leu Arg Asn Thr Thr Thr Gly Glu Glu Thr Thr Leu Val Val Thr Gly
      225     230     235
Val Phe Val Ala Ile Gly His Glu Pro Arg Ser Ser Leu Val Ser Asp
      245     250     255
Val Val Asp Ile Asp Pro Asp Gly Tyr Val Leu Val Lys Gly Arg Thr
      260     265     270
Thr Ser Thr Ser Met Asp Gly Val Phe Ala Ala Gly Asp Leu Val Asp
      275     280     285
Arg Thr Tyr Arg Gln Ala Ile Thr Ala Ala Gly Ser Gly Cys Ala Ala
      290     295     300
Ala Ile Asp Ala Glu Arg Trp Leu Ala Glu His Ala Gly Ser Lys Ala
      305     310     315
Asn Glu Thr Thr Glu Glu Thr Gly Asp Val Asp Ser Thr Asp Thr Thr
      325     330     335
Asp Trp Ser Thr Ala Met Thr Asp Ala Lys Asn Ala Gly Val Thr Ile
      340     345     350
Glu Val Thr Asp Ala Ser Phe Phe Ala Asp Val Leu Ser Ser Asn Lys
      355     360     365
Pro Val Leu Val Asp Phe Trp Ala Thr Trp Cys Gly Pro Cys Lys Met
      370     375     380
Val Ala Pro Val Leu Glu Glu Ile Ala Ser Glu Gln Arg Asn Gln Leu
      385     390     395
Thr Val Ala Lys Leu Asp Val Asp Thr Asn Pro Glu Met Ala Arg Glu
      405     410     415
Phe Gln Val Val Ser Ile Pro Thr Met Ile Leu Phe Gln Gly Gly Gln
      420     425     430
Pro Val Lys Arg Ile Val Gly Ala Lys Gly Lys Ala Ala Leu Leu Arg
      435     440     445
Asp Leu Ser Asp Val Val Pro Asn Leu Asn
      450     455

```

<210> 215

<211> 315

<212> PRT

<213> Mycoplasma pneumoniae

<400> 215

```

Met Leu Lys Val Lys Ser Asp Phe Leu Thr Lys Asp Gln Val Ile Tyr
 1      5      10      15
Asp Val Ala Ile Val Gly Ala Gly Pro Ala Gly Ile Ala Ala Gly Ile

```

```
<210> 216
<211> 311
<212> PRT
<213> Mycobacterium smegmatis
```

<400>	216															
Met	Ser	Thr	Ser	Gln	Thr	Val	His	Asp	Val	Ile	Ile	Ile	Gly	Ser	Gly	
1				5					10					15		
Pro	Ala	Gly	Tyr	Thr	Ala	Ala	Ile	Tyr	Ala	Ala	Arg	Ala	Gln	Leu	Lys	
			20					25					30			
Pro	Leu	Val	Phe	Glu	Gly	Thr	Gln	Phe	Gly	Gly	Ala	Leu	Met	Thr	Thr	
		35					40					45				
Thr	Glu	Val	Glu	Asn	Tyr	Pro	Gly	Phe	Arg	Glu	Gly	Ile	Thr	Gly	Pro	
		50				55					60					
Glu	Leu	Met	Asp	Gln	Met	Arg	Glu	Gln	Ala	Leu	Arg	Phe	Arg	Ala	Asp	
65					70					75					80	
Leu	Arg	Met	Glu	Asp	Val	Asp	Ala	Val	Gln	Leu	Glu	Gly	Pro	Val	Lys	
				85					90					95		
Thr	Val	Val	Val	Gly	Asp	Glu	Thr	His	Gln	Ala	Arg	Ala	Val	Ile	Leu	
			100					105					110			
Ala	Met	Gly	Ala	Ala	Ala	Arg	His	Leu	Gly	Val	Pro	Gly	Glu	Glu	Ala	
		115					120					125				
Leu	Thr	Gly	Met	Gly	Val	Ser	Thr	Cys	Ala	Thr	Cys	Asp	Gly	Phe	Phe	
		130				135						140				
Phe	Arg	Asp	Gln	Asp	Ile	Val	Val	Val	Gly	Gly	Gly	Asp	Ser	Ala	Met	
145					150					155					160	
Glu	Glu	Ala	Thr	Phe	Leu	Thr	Arg	Phe	Ala	Arg	Ser	Val	Thr	Leu	Ile	
				165					170					175		

His Arg Arg Asp Glu Phe Arg Ala Ser Lys Ile Met Leu Glu Arg Ala
 180 185 190
 Arg Ala Asn Glu Lys Ile Thr Phe Leu Thr Asn Thr Glu Ile Thr Gln
 195 200 205
 Ile Glu Gly Asp Pro Lys Val Thr Gly Val Arg Leu Arg Asp Thr Val
 210 215 220
 Thr Gly Glu Glu Ser Lys Leu Asp Val Thr Gly Val Phe Val Ala Ile
 225 230 235 240
 Gly His Asp Pro Arg Ser Glu Leu Val Arg Gly Gln Val Glu Leu Asp
 245 250 255
 Asp Glu Gly Tyr Val Lys Val Gln Gly Arg Thr Thr Tyr Thr Ser Leu
 260 265 270
 Asp Gly Val Phe Ala Ala Gly Asp Leu Val Asp His Thr Tyr Arg Gln
 275 280 285
 Ala Ile Thr Ala Ala Gly Ser Gly Cys Ala Ala Ser Ile Asp Ala Glu
 290 295 300
 Arg Trp Leu Ala Glu Gln Asp
 305 310

<210> 217

<211> 335

<212> PRT

<213> Mycobacterium tuberculosis

<400> 217

Met Thr Ala Pro Pro Val His Asp Arg Ala His His Pro Val Arg Asp
 1 5 10 15
 Val Ile Val Ile Gly Ser Gly Pro Ala Gly Tyr Thr Ala Ala Leu Tyr
 20 25 30
 Ala Ala Arg Ala Gln Leu Ala Pro Leu Val Phe Glu Gly Thr Ser Phe
 35 40 45
 Gly Gly Ala Leu Met Thr Thr Thr Asp Val Glu Asn Tyr Pro Gly Phe
 50 55 60
 Arg Asn Gly Ile Thr Gly Pro Glu Leu Met Asp Glu Met Arg Glu Gln
 65 70 75 80
 Ala Leu Arg Phe Gly Ala Asp Leu Arg Met Glu Asp Val Glu Ser Val
 85 90 95
 Ser Leu His Gly Pro Leu Lys Ser Val Val Thr Ala Asp Gly Gln Thr
 100 105 110
 His Arg Ala Arg Ala Val Ile Leu Ala Met Gly Ala Ala Arg Tyr
 115 120 125
 Leu Gln Val Pro Gly Glu Gln Glu Leu Leu Gly Arg Gly Val Ser Ser
 130 135 140
 Cys Ala Thr Cys Asp Gly Phe Phe Phe Arg Asp Gln Asp Ile Ala Val
 145 150 155 160
 Ile Gly Gly Gly Asp Ser Ala Met Glu Glu Ala Thr Phe Leu Thr Arg
 165 170 175
 Phe Ala Arg Ser Val Thr Leu Val His Arg Arg Asp Glu Phe Arg Ala
 180 185 190
 Ser Lys Ile Met Leu Asp Arg Ala Arg Asn Asn Asp Lys Ile Arg Phe
 195 200 205
 Leu Thr Asn His Thr Val Val Ala Val Asp Gly Asp Thr Thr Val Thr
 210 215 220
 Gly Leu Arg Val Arg Asp Thr Asn Thr Gly Ala Glu Thr Thr Leu Pro
 225 230 235 240
 Val Thr Gly Val Phe Val Ala Ile Gly His Glu Pro Arg Ser Gly Leu
 245 250 255
 Val Arg Glu Ala Ile Asp Val Asp Pro Asp Gly Tyr Val Leu Val Gln
 260 265 270
 Gly Arg Thr Thr Ser Thr Ser Leu Pro Gly Val Phe Ala Ala Gly Asp
 275 280 285
 Leu Val Asp Arg Thr Tyr Arg Gln Ala Val Thr Ala Ala Gly Ser Gly
 290 295 300
 Cys Ala Ala Ala Ile Asp Ala Glu Arg Trp Leu Ala Glu His Ala Ala
 305 310 315 320
 Thr Gly Glu Ala Asp Ser Thr Asp Ala Leu Ile Gly Ala Gln Arg

325

330

335

<210> 218
 <211> 334
 <212> PRT
 <213> *Neurospora crassa*

<400> 218
 Met His Ser Lys Val Val Ile Ile Gly Ser Gly Pro Ala Ala His Thr
 1 5 10 15
 Ala Ala Ile Tyr Leu Ala Arg Ala Glu Leu Lys Pro Val Leu Tyr Glu
 20 25 30
 Gly Phe Met Ala Asn Gly Ile Ala Ala Gly Gly Gln Leu Thr Thr Thr
 35 40 45
 Thr Glu Ile Glu Asn Phe Pro Gly Phe Pro Asp Gly Ile Met Gly Gln
 50 55 60
 Glu Leu Met Asp Lys Met Lys Ala Gln Ser Glu Arg Phe Gly Thr Gln
 65 70 75 80
 Ile Ile Ser Glu Thr Val Ala Lys Val Asp Leu Ser Ala Arg Pro Phe
 85 90 95
 Lys Tyr Ala Thr Glu Trp Ser Pro Glu Glu Tyr His Thr Ala Asp Ser
 100 105 110
 Ile Ile Leu Ala Thr Gly Ala Ser Ala Arg Arg Leu His Leu Pro Gly
 115 120 125
 Glu Glu Lys Tyr Trp Gln Asn Gly Ile Ser Ala Cys Ala Val Cys Asp
 130 135 140
 Gly Ala Val Pro Ile Phe Arg Asn Lys His Leu Val Val Ile Gly Gly
 145 150 155 160
 Gly Asp Ser Ala Ala Glu Glu Ala Met Tyr Leu Thr Lys Tyr Gly Ser
 165 170 175
 His Val Thr Val Leu Val Arg Lys Asp Lys Leu Arg Ala Ser Ser Ile
 180 185 190
 Met Ala His Arg Leu Leu Asn His Glu Lys Val Thr Val Arg Phe Asn
 195 200 205
 Thr Val Gly Val Glu Val Lys Gly Asp Asp Lys Gly Leu Met Ser His
 210 215 220
 Leu Val Val Lys Asp Val Thr Thr Gly Lys Glu Glu Thr Leu Glu Ala
 225 230 235 240
 Asn Gly Leu Phe Tyr Ala Ile Gly His Asp Pro Ala Thr Ala Leu Val
 245 250 255
 Lys Gly Gln Leu Glu Thr Asp Ala Asp Gly Tyr Val Val Thr Lys Pro
 260 265 270
 Gly Thr Thr Leu Thr Ser Val Glu Gly Val Phe Ala Ala Gly Asp Val
 275 280 285
 Gln Asp Lys Arg Tyr Arg Gln Ala Ile Thr Ser Ala Gly Thr Gly Cys
 290 295 300
 Met Ala Ala Leu Asp Ala Glu Lys Phe Leu Ser Glu His Glu Glu Thr
 305 310 315 320
 Pro Ala Glu His Arg Asp Thr Ser Ala Val Gln Gly Asn Leu
 325 330

<210> 219
 <211> 333
 <212> PRT
 <213> *Penicillium chrysogenum*

<400> 219
 Val His Ser Lys Val Val Ile Ile Gly Ser Gly Ala Gly Ala His Thr
 1 5 10 15
 Ala Ala Ile Tyr Leu Ser Arg Ala Glu Leu Gln Pro Val Leu Tyr Glu
 20 25 30
 Gly Met Leu Ala Asn Gly Thr Ala Ala Gly Gly Gln Leu Thr Thr Thr
 35 40 45
 Thr Asp Val Glu Asn Phe Pro Gly Phe Pro Ser Gly Ile Gly Gly Ala
 50 55 60

Glu Leu Met Asp Asn Met Arg Ala Gln Ser Glu Arg Phe Gly Thr Glu
 65 70 75 80
 Ile Ile Thr Glu Thr Ile Ser Lys Leu Asp Leu Ser Ser Arg Pro Phe
 85 90 95
 Lys Met Trp Thr Glu Trp Asn Asp Asp Glu Gly Ser Glu Pro Val Arg
 100 105 110
 Thr Ala Asp Ala Val Ile Ile Ala Thr Gly Ala Asn Ala Arg Arg Leu
 115 120 125
 Asn Leu Pro Gly Glu Glu Thr Tyr Trp Gln Asn Gly Ile Ser Ala Cys
 130 135 140
 Ala Val Cys Asp Gly Ala Val Pro Ile Phe Arg Asn Lys Pro Leu Tyr
 145 150 155 160
 Val Ile Gly Gly Gly Asp Ser Ala Ala Glu Glu Ala Met Phe Leu Ala
 165 170 175
 Lys Tyr Gly Ser Ser Val Thr Val Leu Val Arg Lys Asp Lys Leu Arg
 180 185 190
 Ala Ser Asn Ile Met Ala Asp Arg Leu Leu Ala His Pro Lys Cys Lys
 195 200 205
 Val Arg Phe Asn Thr Val Ala Thr Glu Val Ile Gly Glu Asn Lys Pro
 210 215 220
 Asn Gly Leu Met Thr His Leu Arg Val Lys Asp Val Leu Ser Asn Ala
 225 230 235 240
 Glu Glu Val Val Glu Ala Asn Gly Leu Phe Tyr Ala Val Gly His Asp
 245 250 255
 Pro Ala Ser Gly Leu Val Lys Gly Gln Val Glu Leu Asp Asp Glu Gly
 260 265 270
 Tyr Ile Ile Thr Lys Pro Gly Thr Ser Phe Thr Asn Val Glu Gly Val
 275 280 285
 Phe Ala Cys Gly Asp Val Gln Asp Lys Arg Tyr Arg Gln Ala Ile Thr
 290 295 300
 Ser Ala Gly Ser Gly Cys Val Ala Ala Leu Glu Ala Glu Lys Phe Ile
 305 310 315 320
 Ala Glu Thr Glu Thr His Gln Glu Ala Lys Pro Val Leu
 325 330

<210> 220

<211> 310

<212> PRT

<213> Rickettsia prowazekii

<400> 220

Met Lys Ile Thr Thr Lys Val Leu Ile Ile Gly Ser Gly Pro Ala Gly
 1 5 10 15
 Leu Ser Ala Ala Ile Tyr Thr Ala Arg Ser Ala Leu Lys Pro Ile Leu
 20 25 30
 Ile Asn Gly Met Gln Pro Gly Gly Gln Leu Thr Met Thr Thr Asp Val
 35 40 45
 Glu Asn Tyr Pro Gly Phe Ala Glu Thr Ile Gln Gly Pro Trp Leu Met
 50 55 60
 Glu Gln Met Ser Met Gln Ala Lys Asn Val Gly Thr Glu Ile Ile Ser
 65 70 75 80
 Asp Tyr Val Glu Arg Val Asp Leu Ser Lys Arg Pro Phe Lys Ile Phe
 85 90 95
 Thr Gly Thr Gly Asn Glu Tyr Glu Ala Asp Ser Ile Ile Ile Cys Thr
 100 105 110
 Gly Ala Glu Ser Lys Trp Leu Gly Ile Ala Ser Glu Gln Glu Phe Arg
 115 120 125
 Gly Phe Gly Val Ser Ser Cys Ala Ile Cys Asp Gly Phe Phe Phe Lys
 130 135 140
 Asn Gln Glu Ile Val Val Val Gly Gly Gly Asn Ser Ala Leu Glu Glu
 145 150 155 160
 Ala Leu Tyr Leu Thr Asn His Ala Asn Lys Val Thr Val Val His Arg
 165 170 175
 Arg Asn Ser Phe Arg Ala Glu Lys Ile Leu Gln Asp Arg Leu Phe Lys
 180 185 190
 Asn Pro Lys Ile Ser Val Ile Trp Asp His Ile Ile Asp Glu Ile Val

```

      195      200      205
Gly Ser Asn Lys Pro Lys Ala Val Thr Gly Val Lys Ile Gln Asn Val
  210      215      220
Tyr Thr Asn Glu Ile Asn Leu Val Asn Cys Ser Gly Val Phe Ile Ala
  225      230      235
Ile Gly His Ala Pro Asn Thr Ala Leu Phe Lys Gly Gln Ile Ala Ile
      245      250      255
Asp Asp Asp Asn Tyr Ile Val Thr Gln Ser Gly Ser Thr Arg Thr Asn
      260      265      270
Val Glu Gly Val Phe Ala Ala Gly Asp Val Gln Asp Lys Ile Tyr Arg
      275      280      285
Gln Ala Val Thr Ala Ala Ala Ser Gly Cys Met Ala Ala Leu Glu Val
      290      295      300
Ala Lys Phe Leu Asn Lys
  305      310

```

<210> 221

<211> 322

<212> PRT

<213> Schizosaccharomyces pombe

```

<400> 221
Met Thr His Asn Lys Val Val Ile Ile Gly Ser Gly Pro Ala Gly His
  1      5      10      15
Thr Ala Ala Ile Tyr Leu Ala Arg Gly Glu Leu Lys Pro Val Met Tyr
      20      25      30
Glu Gly Met Leu Ala Asn Gly Ile Ala Ala Gly Gly Gln Leu Thr Thr
      35      40      45
Thr Thr Asp Val Glu Asn Phe Pro Gly Phe Pro Asp Gly Ile Asn Gly
      50      55      60
Thr Thr Leu Thr Glu Asn Phe Arg Ala Gln Ser Leu Arg Phe Gly Thr
      65      70      75      80
Glu Ile Ile Thr Glu Thr Val Ser Lys Leu Asp Leu Ser Ser Arg Pro
      85      90      95
Phe Lys Tyr Trp Leu Glu Gly Ala Glu Glu Glu Glu Pro His Thr Ala
      100      105      110
Asp Ser Val Ile Leu Ala Thr Gly Ala Ser Ala Arg Arg Leu His Ile
      115      120      125
Thr Gly Glu Asp Thr Tyr Trp Gln Ala Gly Ile Ser Ala Cys Ala Val
      130      135      140
Cys Asp Gly Ala Val Pro Ile Tyr Arg Asn Lys Pro Leu Ala Val Val
      145      150      155      160
Gly Gly Gly Asp Ser Ala Ala Glu Glu Ala Gln Phe Leu Thr Lys Tyr
      165      170      175
Gly Ser Lys Val Tyr Val Leu Val Arg Arg Asp Lys Leu Arg Ala Ser
      180      185      190
Pro Ile Met Ala Lys Arg Leu Leu Ala Asn Pro Lys Val Glu Val Leu
      195      200      205
Trp Asn Thr Val Ala Glu Glu Ala Gln Gly Asp Gly Lys Leu Leu Asn
      210      215      220
Asn Leu Arg Ile Lys Asn Thr Asn Thr Asn Glu Val Ser Asp Leu Gln
      225      230      235
Val Asn Gly Leu Phe Tyr Ala Ile Gly His Ile Pro Ala Thr Lys Leu
      245      250      255
Val Ala Glu Gln Ile Glu Leu Asp Glu Ala Gly Tyr Ile Lys Thr Ile
      260      265      270
Asn Gly Thr Pro Arg Thr Ser Ile Pro Gly Phe Phe Ala Ala Gly Asp
      275      280      285
Val Gln Asp Lys Val Phe Arg Gln Ala Ile Thr Ser Ala Gly Ser Gly
      290      295      300
Cys Gln Ala Ala Leu Leu Ala Met His Tyr Leu Glu Glu Leu Glu Asp
      305      310      315      320
Thr Asp

```

<210> 222
 <211> 321
 <212> PRT
 <213> Streptomyces clavuligerus

<400> 222
 Ser Asp Val Arg Asn Val Ile Ile Ile Gly Ser Gly Pro Ala Gly Tyr
 1 5 10 15
 Thr Ala Ala Leu Tyr Thr Ala Arg Ala Ser Leu Gln Pro Leu Val Phe
 20 25 30
 Glu Gly Ala Val Thr Ala Gly Gly Ala Leu Met Asn Thr Asp Val
 35 40 45
 Glu Asn Phe Pro Gly Phe Arg Asp Gly Ile Met Gly Pro Asp Leu Met
 50 55 60
 Asp Asn Met Arg Ala Gln Ala Glu Arg Phe Gly Ala Glu Leu Ile Pro
 65 70 75 80
 Asp Asp Val Val Ser Val Asp Leu Thr Gly Asp Ile Lys Thr Val Thr
 85 90 95
 Asp Ser Ala Gly Thr Val His Arg Ala Lys Ala Val Ile Val Thr Thr
 100 105 110
 Gly Ser Gln His Arg Lys Leu Gly Leu Pro Arg Glu Asp Ala Leu Ser
 115 120 125
 Gly Arg Gly Val Ser Trp Cys Ala Thr Cys Asp Gly Phe Phe Phe Lys
 130 135 140
 Asp Gln Asp Ile Val Val Val Gly Gly Gly Asp Thr Ala Met Glu Glu
 145 150 155 160
 Ala Thr Phe Leu Ser Arg Phe Ala Lys Ser Val Thr Ile Val His Arg
 165 170 175
 Arg Asp Ser Leu Arg Ala Ser Lys Ala Met Gln Asp Arg Ala Phe Ala
 180 185 190
 Asp Pro Lys Ile Ser Phe Ala Trp Asn Ser Glu Val Ala Thr Ile His
 195 200 205
 Gly Glu Gln Lys Leu Thr Gly Leu Thr Leu Arg Asp Thr Lys Thr Gly
 210 215 220
 Glu Thr Arg Glu Leu Ala Ala Thr Gly Leu Phe Ile Ala Val Gly His
 225 230 235 240
 Asp Pro Arg Thr Glu Leu Phe Lys Gly Gln Leu Asp Leu Asp Asp Glu
 245 250 255
 Gly Tyr Leu Lys Val Ala Ser Pro Ser Thr Arg Thr Asn Leu Thr Gly
 260 265 270
 Val Phe Ala Ala Gly Asp Val Val Asp His Thr Tyr Arg Gln Ala Ile
 275 280 285
 Thr Ala Ala Gly Thr Gly Cys Ser Ala Ala Leu Asp Ala Glu Arg Tyr
 290 295 300
 Leu Ala Ala Leu Ala Asp Ser Glu Gln Ile Ala Glu Pro Ala Pro Ala
 305 310 315 320
 Val

<210> 223
 <211> 321
 <212> PRT
 <213> Streptomyces coelicolor

<400> 223
 Ser Asp Val Arg Asn Val Ile Ile Ile Gly Ser Gly Pro Ala Gly Tyr
 1 5 10 15
 Thr Ala Ala Leu Tyr Thr Ala Arg Ala Ser Leu Lys Pro Leu Val Phe
 20 25 30
 Glu Gly Ala Val Thr Ala Gly Gly Ala Leu Met Asn Thr Thr Glu Val
 35 40 45
 Glu Asn Phe Pro Gly Phe Gln Asp Gly Ile Met Gly Pro Glu Leu Met
 50 55 60
 Asp Asn Met Arg Ala Gln Ala Glu Arg Phe Gly Ala Glu Leu Ile Pro
 65 70 75 80
 Asp Asp Val Val Ala Val Asp Leu Ser Gly Glu Ile Lys Thr Val Thr

Asp	Thr	Ala	Gly	85	Thr	Val	His	Arg	Ala	Lys	90	Ala	Val	Ile	Val	Thr	Thr	95
			100							105					110			
Gly	Ser	Gln	His	Arg	Lys	Leu	Gly	Leu	Pro	Asn	Glu	Asp	Ala	Leu	Ser			
		115					120					125						
Gly	Arg	Gly	Val	Ser	Trp	Cys	Ala	Thr	Cys	Asp	Gly	Phe	Phe	Phe	Lys			
	130					135					140							
Asp	Gln	Asp	Ile	Ala	Val	Ile	Gly	Gly	Gly	Asp	Thr	Ala	Met	Glu	Glu			
145					150					155					160			
Ala	Thr	Phe	Leu	Ser	Arg	Phe	Ala	Lys	Ser	Val	Thr	Ile	Val	His	Arg			
			165						170					175				
Arg	Asp	Thr	Leu	Arg	Ala	Ser	Lys	Ala	Met	Gln	Glu	Arg	Ala	Phe	Ala			
		180							185					190				
Asp	Pro	Lys	Ile	Ser	Phe	Val	Trp	Asp	Ser	Glu	Val	Ala	Glu	Val	Gln			
	195						200					205						
Gly	Asp	Gln	Lys	Leu	Ala	Gly	Leu	Lys	Leu	Arg	Asn	Val	Lys	Thr	Gly			
	210					215					220							
Glu	Leu	Ser	Asp	Leu	Pro	Val	Thr	Gly	Leu	Phe	Ile	Ala	Ile	Gly	His			
225					230					235					240			
Asp	Pro	Arg	Thr	Glu	Leu	Phe	Lys	Gly	Gln	Leu	Asp	Leu	Asp	Pro	Glu			
			245						250					255				
Gly	Tyr	Leu	Lys	Val	Asp	Ala	Pro	Ser	Thr	Arg	Thr	Asn	Leu	Thr	Gly			
		260						265					270					
Val	Phe	Gly	Ala	Gly	Asp	Val	Val	Asp	His	Thr	Tyr	Arg	Gln	Ala	Ile			
	275						280					285						
Thr	Ala	Ala	Gly	Thr	Gly	Cys	Ser	Ala	Ala	Val	Asp	Ala	Glu	Pro	Phe			
	290					295					300							
Leu	Ala	Ala	Leu	Ser	Asp	Glu	Asp	Lys	Ala	Glu	Pro	Glu	Lys	Thr	Ala			
305					310					315					320			
Val																		

<210> 224

<211> 307

<212> PRT

<213> Treponema pallidum

<400> 224

Met	Glu	Thr	Asp	Tyr	Asp	Val	Ile	Ile	Val	Gly	Ala	Gly	Ala	Ala	Gly			
1				5					10					15				
Leu	Ser	Ala	Ala	Gln	Tyr	Ala	Cys	Arg	Ala	Asn	Leu	Arg	Thr	Leu	Val			
			20					25					30					
Ile	Glu	Ser	Lys	Ala	His	Gly	Gly	Gln	Ala	Leu	Leu	Ile	Asp	Ser	Leu			
	35					40						45						
Glu	Asn	Tyr	Pro	Gly	Tyr	Ala	Thr	Pro	Ile	Ser	Gly	Phe	Glu	Tyr	Ala			
	50					55					60							
Glu	Asn	Met	Lys	Lys	Gln	Ala	Val	Ala	Phe	Gly	Ala	Gln	Ile	Ala	Tyr			
65					70					75					80			
Glu	Glu	Val	Thr	Thr	Ile	Gly	Lys	Arg	Asp	Ser	Val	Phe	His	Ile	Thr			
			85						90					95				
Thr	Gly	Thr	Gly	Ala	Tyr	Thr	Ala	Met	Ser	Val	Ile	Leu	Ala	Thr	Gly			
		100						105					110					
Ala	Glu	His	Arg	Lys	Met	Gly	Ile	Pro	Gly	Glu	Ser	Glu	Phe	Leu	Gly			
		115				120						125						
Arg	Gly	Val	Ser	Tyr	Cys	Ala	Thr	Cys	Asp	Gly	Pro	Phe	Phe	Arg	Asn			
	130					135					140							
Lys	His	Val	Val	Val	Ile	Gly	Gly	Gly	Asp	Ala	Ala	Cys	Asp	Glu	Ser			
145					150					155					160			
Leu	Val	Leu	Ser	Arg	Leu	Thr	Asp	Arg	Val	Thr	Met	Ile	His	Arg	Arg			
			165						170					175				
Asp	Thr	Leu	Arg	Ala	Gln	Lys	Ala	Ile	Ala	Glu	Arg	Thr	Leu	Lys	Asn			
		180						185					190					
Pro	His	Ile	Ala	Val	Gln	Trp	Asn	Thr	Thr	Leu	Glu	Ala	Val	Arg	Gly			
	195					200						205						
Glu	Thr	Lys	Val	Ser	Ser	Val	Leu	Leu	Lys	Asp	Val	Lys	Thr	Gly	Glu			
	210					215					220							

Thr Arg Glu Leu Ala Cys Asp Ala Val Phe Phe Phe Ile Gly Met Val
 225 230 235 240
 Pro Ile Thr Gly Leu Leu Pro Asp Ala Glu Lys Asp Ser Thr Gly Tyr
 245 250 255
 Ile Val Thr Asp Asp Glu Met Arg Thr Ser Val Glu Gly Ile Phe Ala
 260 265 270
 Ala Gly Asp Val Arg Ala Lys Ser Phe Arg Gln Val Ile Thr Ala Thr
 275 280 285
 Ser Asp Gly Ala Leu Ala Ala His Ala Ala Ala Ser Tyr Ile Asp Thr
 290 295 300
 Leu Gln Asn
 305

<210> 225

<211> 45

<212> PRT

<213> *Vibrio fischeri*

<400> 225

Met Asn Val Lys His Ser Lys Leu Leu Ile Leu Gly Ser Gly Pro Ala
 1 5 10 15
 Gly Tyr Thr Ala Ala Val Tyr Ala Ala Arg Ala Asn Leu Asn Pro Val
 20 25 30
 Met Ile Thr Gly Met Gln Gln Gly Gly Gln Leu Thr Asn
 35 40 45

<210> 226

<211> 318

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 226

Val His Asn Lys Val Thr Ile Ile Gly Ser Gly Pro Ala Ala His Thr
 1 5 10 15
 Ala Ala Ile Tyr Leu Ala Arg Ala Glu Ile Lys Pro Ile Leu Tyr Glu
 20 25 30
 Gly Met Met Ala Asn Gly Ile Ala Ala Gly Gly Gln Leu Thr Thr Thr
 35 40 45
 Thr Glu Ile Glu Asn Phe Pro Gly Phe Pro Asp Gly Leu Thr Gly Ser
 50 55 60
 Glu Leu Met Asp Arg Met Arg Glu Gln Ser Thr Lys Phe Gly Thr Glu
 65 70 75 80
 Ile Ile Thr Glu Thr Val Ser Lys Val Asp Leu Ser Ser Lys Pro Phe
 85 90 95
 Lys Leu Trp Thr Glu Phe Asn Glu Asp Ala Glu Pro Val Thr Thr Asp
 100 105 110
 Ala Ile Ile Leu Ala Thr Gly Ala Ser Ala Lys Arg Met His Leu Pro
 115 120 125
 Gly Glu Glu Thr Tyr Trp Gln Lys Gly Ile Ser Ala Cys Ala Val Cys
 130 135 140
 Asp Gly Ala Val Pro Ile Phe Arg Asn Lys Pro Leu Ala Val Ile Gly
 145 150 155 160
 Gly Gly Asp Ser Ala Cys Glu Glu Ala Gln Phe Leu Thr Lys Tyr Gly
 165 170 175
 Ser Lys Val Phe Met Leu Val Arg Lys Asp His Leu Arg Ala Ser Thr
 180 185 190
 Ile Met Gln Lys Arg Ala Glu Lys Asn Glu Lys Ile Glu Ile Leu Tyr
 195 200 205
 Asn Thr Val Ala Leu Glu Ala Lys Gly Asp Gly Lys Leu Leu Asn Ala
 210 215 220
 Leu Arg Ile Lys Asn Thr Lys Lys Asn Glu Glu Thr Asp Leu Pro Val
 225 230 235 240
 Ser Gly Leu Phe Tyr Ala Ile Gly His Thr Pro Ala Thr Lys Ile Val
 245 250 255
 Ala Gly Gln Val Asp Thr Asp Glu Ala Gly Tyr Ile Lys Thr Val Pro

			260					265				270			
Gly	Ser	Ser	Leu	Thr	Ser	Val	Pro	Gly	Phe	Phe	Ala	Ala	Gly	Asp	Val
			275				280					285			
Gln	Asp	Ser	Lys	Tyr	Arg	Gln	Ala	Ile	Thr	Ser	Ala	Gly	Ser	Gly	Cys
			290			295					300				
Met	Ala	Ala	Leu	Asp	Ala	Glu	Lys	Tyr	Leu	Thr	Ser	Leu	Glu		
305					310					315					

```
<210> 227
<211> 342
<212> PRT
<213> Saccharomyces cerevisiae
```

<400>	227															
Met	Ile	Lys	His	Ile	Val	Ser	Pro	Phe	Arg	Thr	Asn	Phe	Val	Gly	Ile	
1				5				10						15		
Ser	Lys	Ser	Val	Leu	Ser	Arg	Met	Ile	His	His	Lys	Val	Thr	Ile	Ile	
			20					25					30			
Gly	Ser	Gly	Pro	Ala	Ala	His	Thr	Ala	Ala	Ile	Tyr	Leu	Ala	Arg	Ala	
		35					40					45				
Glu	Met	Lys	Pro	Thr	Leu	Tyr	Glu	Gly	Met	Met	Ala	Asn	Gly	Ile	Ala	
50						55					60					
Ala	Gly	Gly	Gln	Leu	Thr	Thr	Thr	Thr	Asp	Ile	Glu	Asn	Phe	Pro	Gly	
65				70					75						80	
Phe	Pro	Glu	Ser	Leu	Ser	Gly	Ser	Glu	Leu	Met	Glu	Arg	Met	Arg	Lys	
			85					90						95		
Gln	Ser	Ala	Lys	Phe	Gly	Thr	Asn	Ile	Ile	Thr	Glu	Thr	Val	Ser	Lys	
			100					105					110			
Val	Asp	Leu	Ser	Ser	Lys	Pro	Phe	Arg	Leu	Trp	Thr	Glu	Phe	Asn	Glu	
		115					120					125				
Asp	Ala	Glu	Pro	Val	Thr	Thr	Asp	Ala	Ile	Ile	Leu	Ala	Thr	Gly	Ala	
		130				135					140					
Ser	Ala	Lys	Arg	Met	His	Leu	Pro	Gly	Glu	Glu	Thr	Tyr	Trp	Gln	Gln	
145					150				155						160	
Gly	Ile	Ser	Ala	Cys	Ala	Val	Cys	Asp	Gly	Ala	Val	Pro	Ile	Phe	Arg	
			165					170						175		
Asn	Lys	Pro	Leu	Ala	Val	Ile	Gly	Gly	Gly	Asp	Ser	Ala	Cys	Glu	Glu	
			180					185					190			
Ala	Glu	Phe	Leu	Thr	Lys	Tyr	Ala	Ser	Lys	Val	Tyr	Ile	Leu	Val	Arg	
		195					200					205				
Lys	Asp	His	Phe	Arg	Ala	Ser	Val	Ile	Met	Gln	Arg	Arg	Ile	Glu	Lys	
	210					215					220					
Asn	Pro	Asn	Ile	Ile	Val	Leu	Phe	Asn	Thr	Val	Ala	Leu	Glu	Ala	Lys	
225					230					235					240	
Gly	Asp	Gly	Lys	Leu	Leu	Asn	Met	Leu	Arg	Ile	Lys	Asn	Thr	Lys	Ser	
			245					250						255		
Asn	Val	Glu	Asn	Asp	Leu	Glu	Val	Asn	Gly	Leu	Phe	Tyr	Ala	Ile	Gly	
			260					265					270			
His	Ser	Pro	Ala	Thr	Asp	Ile	Val	Lys	Gly	Gln	Val	Asp	Glu	Glu	Glu	
		275					280					285				
Thr	Gly	Tyr	Ile	Lys	Thr	Val	Pro	Gly	Ser	Ser	Leu	Thr	Ser	Val	Pro	
	290					295					300					
Gly	Phe	Phe	Ala	Ala	Gly	Asp	Val	Gln	Asp	Ser	Arg	Tyr	Arg	Gln	Ala	
305					310				315						320	
Val	Thr	Ser	Ala	Gly	Ser	Gly	Cys	Ile	Ala	Ala	Leu	Asp	Ala	Glu	Arg	
			325					330								

```
<210> 228
<211> 499
<212> PRT
<213> Bos taurus
```

<400> 228

Met Asn Gly Ser Lys Asp Leu Pro Glu Pro Tyr Asp Tyr Asp Leu Ile
 1 5 10 15
 Ile Ile Gly Gly Gly Ser Gly Gly Leu Ala Ala Ala Lys Glu Ala Ala
 20 25 30
 Lys Tyr Asp Lys Lys Val Met Val Leu Asp Phe Val Thr Pro Thr Pro
 35 40 45
 Leu Gly Thr Arg Trp Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys
 50 55 60
 Ile Pro Lys Lys Leu Met His Gln Ala Ala Leu Leu Gly Gln Ala Leu
 65 70 75 80
 Arg Asp Ser Arg Asn Tyr Gly Trp Asn Val Glu Glu Thr Val Lys His
 85 90 95
 Asp Trp Glu Arg Met Thr Glu Ala Val Gln Asn His Ile Gly Ser Leu
 100 105 110
 Asn Trp Gly Tyr Arg Val Ala Leu Arg Glu Lys Lys Val Thr Tyr Glu
 115 120 125
 Asn Ala Tyr Gly Glu Phe Val Gly Pro His Arg Ile Lys Ala Thr Asn
 130 135 140
 Asn Lys Gly Lys Glu Lys Ile Tyr Ser Ala Glu Arg Phe Leu Ile Ala
 145 150 155 160
 Thr Gly Glu Arg Pro Arg Tyr Leu Gly Ile Pro Gly Asp Lys Glu Tyr
 165 170 175
 Cys Ile Ser Ser Asp Asp Leu Phe Ser Leu Pro Tyr Cys Pro Gly Lys
 180 185 190
 Thr Leu Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe
 195 200 205
 Leu Ala Gly Ile Gly Leu Asp Val Thr Val Met Val Arg Ser Ile Leu
 210 215 220
 Leu Arg Gly Phe Asp Gln Asp Met Ala Asn Lys Ile Gly Glu His Met
 225 230 235 240
 Gln Glu His Gly Ile Lys Phe Ile Arg Gln Phe Val Pro Ile Lys Val
 245 250 255
 Glu Gln Ile Glu Ala Gly Thr Pro Gly Arg Leu Arg Val Ile Ala Lys
 260 265 270
 Ser Thr Asp Ser Asp Gln Thr Ile Glu Gly Glu Tyr Asn Thr Val Leu
 275 280 285
 Leu Ala Ile Gly Arg Asp Ala Cys Thr Arg Lys Ile Gly Leu Glu Asn
 290 295 300
 Val Gly Val Lys Ile Asn Glu Lys Thr Gly Lys Ile Pro Val Thr Glu
 305 310 315 320
 Glu Glu Gln Thr Asn Val Pro Tyr Ile Tyr Ala Ile Gly Asp Ile Leu
 325 330 335
 Glu Gly Lys Leu Glu Leu Thr Pro Val Ala Ile Gln Ala Gly Arg Leu
 340 345 350
 Leu Ala Gln Arg Leu Tyr Gly Gly Ser Thr Val Lys Cys Asp Tyr Glu
 355 360 365
 Asn Val Pro Thr Thr Val Phe Thr Pro Leu Glu Tyr Gly Ser Cys Gly
 370 375 380
 Leu Ser Glu Glu Lys Ala Val Glu Lys Phe Gly Glu Glu Asn Val Glu
 385 390 395 400
 Val Tyr His Ser Tyr Phe Trp Pro Leu Glu Trp Thr Ile Pro Ser Arg
 405 410 415
 Asp Asn Asn Lys Cys Tyr Ala Lys Val Val Cys Asn Ile Lys Asp Asn
 420 425 430
 Glu Arg Val Val Gly Phe His Val Leu Gly Pro Asn Ala Gly Glu Val
 435 440 445
 Thr Gln Gly Phe Ala Ala Ala Leu Lys Cys Gly Leu Thr Lys Asp Gln
 450 455 460
 Leu Asp Ser Thr Ile Gly Ile His Pro Val Cys Ala Glu Val Phe Thr
 465 470 475 480
 Thr Leu Ser Val Thr Lys Arg Ser Gly Gly Asn Ile Leu Gln Thr Gly
 485 490 495
 Cys Cys Gly

<210> 229

<211> 523
 <212> PRT
 <213> *Caenorhabditis elegans*

<400> 229
 Met Tyr Ile Lys Gly Asn Ala Val Gly Gly Leu Lys Glu Leu Lys Ala
 1 5 10 15
 Leu Lys Gln Asp Tyr Leu Lys Glu Trp Leu Arg Asp His Thr Tyr Asp
 20 25 30
 Leu Ile Val Ile Gly Gly Gly Ser Gly Gly Leu Ala Ala Ala Lys Glu
 35 40 45
 Ala Ser Arg Leu Gly Lys Lys Val Ala Cys Leu Asp Phe Val Lys Pro
 50 55 60
 Ser Pro Gln Gly Thr Ser Trp Gly Leu Gly Gly Thr Cys Val Asn Val
 65 70 75 80
 Gly Cys Ile Pro Lys Lys Leu Met His Gln Ala Ser Leu Leu Gly His
 85 90 95
 Ser Ile His Asp Ala Lys Lys Tyr Gly Trp Lys Leu Pro Glu Gly Lys
 100 105 110
 Val Glu His Gln Trp Asn His Leu Arg Asp Ser Val Gln Asp His Ile
 115 120 125
 Ala Ser Leu Asn Trp Gly Tyr Arg Val Gln Leu Arg Glu Lys Thr Val
 130 135 140
 Thr Tyr Ile Asn Ser Tyr Gly Glu Phe Thr Gly Pro Phe Glu Ile Ser
 145 150 155 160
 Ala Thr Asn Lys Lys Lys Lys Val Glu Lys Leu Thr Ala Asp Arg Phe
 165 170 175
 Leu Ile Ser Thr Gly Leu Arg Pro Lys Tyr Pro Glu Ile Pro Gly Val
 180 185 190
 Lys Glu Tyr Thr Ile Thr Ser Asp Asp Leu Phe Gln Leu Pro Tyr Ser
 195 200 205
 Pro Gly Lys Thr Leu Cys Val Gly Ala Ser Tyr Val Ser Leu Glu Cys
 210 215 220
 Ala Gly Phe Leu His Gly Phe Gly Phe Asp Val Thr Val Met Val Arg
 225 230 235 240
 Ser Ile Leu Leu Arg Gly Phe Asp Gln Asp Met Ala Glu Arg Ile Arg
 245 250 255
 Lys His Met Ile Ala Tyr Gly Met Lys Phe Glu Ala Gly Val Pro Thr
 260 265 270
 Arg Ile Glu Gln Ile Asp Glu Lys Thr Asp Glu Lys Ala Gly Lys Tyr
 275 280 285
 Arg Val Phe Trp Pro Lys Lys Asn Glu Glu Thr Gly Glu Met Gln Glu
 290 295 300
 Val Ser Glu Glu Tyr Asn Thr Ile Leu Met Ala Ile Gly Arg Glu Ala
 305 310 315 320
 Val Thr Asp Asp Val Gly Leu Thr Thr Ile Gly Val Glu Arg Ala Lys
 325 330 335
 Ser Lys Lys Val Leu Gly Arg Arg Glu Gln Ser Thr Thr Ile Pro Trp
 340 345 350
 Val Tyr Ala Ile Gly Asp Val Leu Glu Gly Thr Pro Glu Leu Thr Pro
 355 360 365
 Val Ala Ile Gln Ala Gly Arg Val Leu Met Arg Arg Ile Phe Asp Gly
 370 375 380
 Ala Asn Glu Leu Thr Glu Tyr Asp Gln Ile Pro Thr Thr Val Phe Thr
 385 390 395 400
 Pro Leu Glu Tyr Gly Cys Cys Gly Leu Ser Glu Glu Asp Ala Met Met
 405 410 415
 Lys Tyr Gly Lys Asp Asn Ile Ile Ile Tyr His Asn Val Phe Asn Pro
 420 425 430
 Leu Glu Tyr Thr Ile Ser Glu Arg Met Asp Lys Asp His Cys Tyr Leu
 435 440 445
 Lys Met Ile Cys Leu Arg Asn Glu Glu Glu Lys Val Val Gly Phe His
 450 455 460
 Ile Leu Thr Pro Asn Ala Gly Glu Val Thr Gln Gly Phe Gly Ile Ala
 465 470 475 480
 Leu Lys Leu Ala Ala Lys Lys Ala Asp Phe Asp Arg Leu Ile Gly Ile
 485 490 495

His Pro Thr Val Ala Glu Asn Phe Thr Thr Leu Thr Leu Glu Lys Lys
 500 505 510
 Glu Gly Asp Glu Glu Leu Gln Ala Ser Gly Cys
 515 520

<210> 230
 <211> 497
 <212> PRT
 <213> Homo sapiens

<400> 230
 Met Asn Gly Pro Glu Asp Leu Pro Lys Ser Tyr Asp Tyr Asp Leu Ile
 1 5 10 15
 Ile Ile Gly Gly Gly Ser Gly Gly Leu Ala Ala Ala Lys Glu Ala Ala
 20 25 30
 Gln Tyr Gly Lys Lys Val Met Val Leu Asp Phe Val Thr Pro Thr Pro
 35 40 45
 Leu Gly Thr Arg Trp Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys
 50 55 60
 Ile Pro Lys Lys Leu Met His Gln Ala Ala Leu Leu Gly Gln Ala Leu
 65 70 75 80
 Gln Asp Ser Arg Asn Tyr Gly Trp Lys Val Glu Glu Thr Val Lys His
 85 90 95
 Asp Trp Asp Arg Met Ile Glu Ala Val Gln Asn His Ile Gly Ser Leu
 100 105 110
 Asn Trp Gly Tyr Arg Val Ala Leu Arg Glu Lys Lys Val Val Tyr Glu
 115 120 125
 Asn Ala Tyr Gly Gln Phe Ile Gly Pro His Arg Ile Lys Ala Thr Asn
 130 135 140
 Asn Lys Gly Lys Glu Lys Ile Tyr Ser Ala Glu Ser Phe Leu Ile Ala
 145 150 155 160
 Thr Gly Glu Arg Pro Arg Tyr Leu Gly Ile Pro Gly Asp Lys Glu Tyr
 165 170 175
 Cys Ile Ser Ser Asp Asp Leu Phe Ser Leu Pro Tyr Cys Pro Gly Lys
 180 185 190
 Thr Leu Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe
 195 200 205
 Leu Ala Gly Ile Gly Leu Gly Val Thr Val Met Val Arg Ser Ile Leu
 210 215 220
 Leu Arg Gly Phe Asp Gln Asp Met Ala Asn Lys Ile Gly Glu His Met
 225 230 235 240
 Glu Glu His Gly Ile Lys Phe Ile Arg Gln Phe Val Pro Ile Lys Val
 245 250 255
 Glu Gln Ile Glu Ala Gly Thr Pro Gly Arg Leu Arg Val Val Ala Gln
 260 265 270
 Ser Thr Asn Ser Glu Glu Ile Ile Glu Gly Glu Tyr Asn Thr Val Met
 275 280 285
 Leu Ala Ile Gly Arg Asp Ala Cys Thr Arg Lys Ile Gly Leu Glu Thr
 290 295 300
 Val Gly Val Lys Ile Asn Glu Lys Thr Gly Lys Ile Pro Val Thr Asp
 305 310 315 320
 Glu Glu Gln Thr Asn Val Pro Tyr Ile Tyr Ala Ile Gly Asp Ile Leu
 325 330 335
 Glu Asp Lys Val Glu Leu Thr Pro Val Ala Ile Gln Ala Gly Arg Leu
 340 345 350
 Leu Ala Gln Arg Leu Tyr Ala Gly Ser Thr Val Lys Cys Asp Tyr Glu
 355 360 365
 Asn Val Pro Thr Thr Val Phe Thr Pro Leu Glu Tyr Gly Ala Cys Gly
 370 375 380
 Leu Ser Glu Glu Lys Ala Val Glu Lys Phe Gly Glu Glu Asn Ile Glu
 385 390 395 400
 Val Tyr His Ser Tyr Phe Trp Pro Leu Glu Trp Thr Ile Pro Ser Arg
 405 410 415
 Asp Asn Asn Lys Cys Tyr Ala Lys Ile Ile Cys Asn Thr Lys Asp Asn
 420 425 430
 Glu Arg Val Val Gly Phe His Val Leu Gly Pro Asn Ala Gly Glu Val

435 440 445
 Thr Gln Gly Phe Ala Ala Leu Lys Cys Gly Leu Thr Lys Lys Gln
 450 455 460
 Leu Asp Ser Thr Ile Gly Ile His Pro Val Cys Ala Glu Val Phe Thr
 465 470 475 480
 Thr Leu Ser Val Thr Lys Arg Ser Gly Ala Ser Ile Leu Gln Ala Gly
 485 490 495
 Cys

<210> 231
 <211> 541
 <212> PRT
 <213> Plasmodium falciparum

<400> 231
 Met Cys Lys Asp Lys Asn Glu Lys Lys Asn Tyr Glu His Val Asn Ala
 1 5 10 15
 Asn Glu Lys Asn Gly Tyr Leu Ala Ser Glu Lys Asn Glu Leu Thr Lys
 20 25 30
 Asn Lys Val Glu Glu His Thr Tyr Asp Tyr Asp Tyr Val Val Ile Gly
 35 40 45
 Gly Gly Pro Gly Gly Met Ala Ser Ala Lys Glu Ala Ala His Gly
 50 55 60
 Ala Arg Val Leu Leu Phe Asp Tyr Val Lys Pro Ser Ser Gln Gly Thr
 65 70 75 80
 Lys Trp Gly Ile Gly Gly Thr Cys Val Asn Val Gly Cys Val Pro Lys
 85 90 95
 Lys Leu Met His Tyr Ala Gly His Met Gly Ser Ile Phe Lys Leu Asp
 100 105 110
 Ser Lys Ala Tyr Gly Trp Lys Phe Asp Asn Leu Lys His Asp Trp Lys
 115 120 125
 Lys Leu Val Thr Thr Val Gln Ser His Ile Arg Ser Leu Asn Phe Ser
 130 135 140
 Tyr Met Thr Gly Leu Arg Ser Ser Lys Val Lys Tyr Ile Asn Gly Leu
 145 150 155 160
 Ala Lys Leu Lys Asp Lys Asn Thr Val Ser Tyr Tyr Leu Lys Gly Asp
 165 170 175
 Leu Ser Lys Glu Glu Thr Val Thr Gly Lys Tyr Ile Leu Ile Ala Thr
 180 185 190
 Gly Cys Arg Pro His Ile Pro Asp Asp Val Glu Gly Ala Lys Glu Leu
 195 200 205
 Ser Ile Thr Ser Asp Asp Ile Phe Ser Leu Lys Lys Asp Pro Gly Lys
 210 215 220
 Thr Leu Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ser Gly Phe
 225 230 235 240
 Leu Asn Ser Leu Gly Tyr Asp Val Thr Val Ala Val Arg Ser Ile Val
 245 250 255
 Leu Arg Gly Phe Asp Gln Gln Cys Ala Val Lys Val Lys Leu Tyr Met
 260 265 270
 Glu Glu Gln Gly Val Met Phe Lys Asn Gly Ile Leu Pro Lys Lys Leu
 275 280 285
 Thr Lys Met Asp Asp Lys Ile Leu Val Glu Phe Ser Asp Lys Thr Ser
 290 295 300
 Glu Leu Tyr Asp Thr Val Leu Tyr Ala Ile Gly Arg Lys Gly Asp Ile
 305 310 315 320
 Asp Gly Leu Asn Leu Glu Ser Leu Asn Met Asn Val Asn Lys Ser Asn
 325 330 335
 Asn Lys Ile Ile Ala Asp His Leu Ser Cys Thr Asn Ile Pro Ser Ile
 340 345 350
 Phe Ala Val Gly Asp Val Ala Glu Asn Val Pro Glu Leu Ala Pro Val
 355 360 365
 Ala Ile Lys Ala Gly Glu Ile Leu Ala Arg Arg Leu Phe Lys Asp Ser
 370 375 380
 Asp Glu Ile Met Asp Tyr Ser Tyr Ile Pro Thr Ser Ile Tyr Thr Pro
 385 390 395 400

Ile Glu Tyr Gly Ala Cys Gly Tyr Ser Glu Glu Lys Ala Tyr Glu Leu
 405 410 415
 Tyr Gly Lys Ser Asn Val Glu Val Phe Leu Gln Glu Phe Asn Asn Leu
 420 425 430
 Glu Ile Ser Ala Val His Arg Gln Lys His Ile Arg Ala Gln Lys Asp
 435 440 445
 Glu Tyr Asp Leu Asp Val Ser Ser Thr Cys Leu Ala Lys Leu Val Cys
 450 455 460
 Leu Lys Asn Glu Asp Asn Arg Val Ile Gly Phe His Tyr Val Gly Pro
 465 470 475 480
 Asn Ala Gly Glu Val Thr Gln Gly Met Ala Leu Ala Leu Arg Leu Lys
 485 490 495
 Val Lys Lys Lys Asp Phe Asp Asn Cys Ile Gly Ile His Pro Thr Asp
 500 505 510
 Ala Glu Ser Phe Met Asn Leu Phe Val Thr Ile Ser Ser Gly Leu Ser
 515 520 525
 Tyr Ala Ala Lys Gly Gly Cys Gly Gly Gly Lys Cys Gly
 530 535 540

<210> 232

<211> 535

<212> PRT

<213> Arabidopsis thaliana

<400> 232

Met Ala Ala Ser Pro Lys Ile Gly Ile Gly Ile Ala Ser Val Ser Ser
 1 5 10 15
 Pro His Arg Val Ser Ala Ala Ser Ser Ala Leu Ser Pro Pro His
 20 25 30
 Leu Phe Phe Leu Thr Thr Thr Thr Thr Arg His Gly Gly Ser Tyr
 35 40 45
 Leu Leu Arg Gln Pro Thr Arg Thr Arg Ser Ser Asp Ser Leu Arg Leu
 50 55 60
 Arg Val Ser Ala Thr Ala Asn Ser Pro Ser Ser Ser Ser Gly Gly
 65 70 75 80
 Glu Ile Ile Glu Asn Val Val Ile Ile Gly Ser Gly Pro Ala Gly Tyr
 85 90 95
 Thr Ala Ala Ile Tyr Ala Ala Arg Ala Asn Leu Lys Pro Val Val Phe
 100 105 110
 Glu Gly Tyr Gln Met Gly Gly Val Pro Gly Gly Gln Leu Met Thr Thr
 115 120 125
 Thr Glu Val Glu Asn Phe Pro Gly Phe Pro Asp Gly Ile Thr Gly Pro
 130 135 140
 Asp Leu Met Glu Lys Met Arg Lys Gln Ala Glu Arg Trp Gly Ala Glu
 145 150 155 160
 Leu Tyr Pro Glu Asp Val Glu Ser Leu Ser Val Thr Thr Ala Pro Phe
 165 170 175
 Thr Val Gln Thr Ser Glu Arg Lys Val Lys Cys His Ser Ile Ile Tyr
 180 185 190
 Ala Thr Gly Ala Thr Ala Arg Arg Leu Arg Leu Pro Arg Glu Glu Glu
 195 200 205
 Phe Trp Ser Arg Gly Ile Ser Ala Cys Ala Ile Cys Asp Gly Ala Ser
 210 215 220
 Pro Leu Phe Lys Gly Gln Val Leu Ala Val Val Gly Gly Gly Asp Thr
 225 230 235 240
 Ala Thr Glu Glu Ala Leu Tyr Leu Thr Lys Tyr Ala Arg His Val His
 245 250 255
 Leu Leu Val Arg Asp Gln Leu Arg Ala Ser Lys Ala Met Gln Asp
 260 265 270
 Arg Val Ile Asn Asn Pro Asn Ile Thr Val His Tyr Asn Thr Glu Thr
 275 280 285
 Val Asp Val Leu Ser Asn Thr Lys Gly Gln Met Ser Gly Ile Leu Leu
 290 295 300
 Arg Arg Leu Asp Thr Gly Glu Glu Thr Glu Leu Glu Ala Lys Gly Leu
 305 310 315 320
 Phe Tyr Gly Ile Gly His Ser Pro Asn Ser Gln Leu Leu Glu Gly Gln

```

          325          330          335
Val Glu Leu Asp Ser Ser Gly Tyr Val Leu Val Arg Glu Gly Thr Ser
          340          345          350
Asn Thr Ser Val Glu Gly Val Phe Ala Ala Gly Asp Val Gln Asp His
          355          360          365
Glu Trp Arg Gln Ala Val Thr Ala Ala Gly Ser Gly Cys Ile Ala Ala
          370          375          380
Leu Ser Ala Glu Arg Tyr Leu Thr Ser Asn Asn Leu Leu Val Glu Phe
          385          390          395          400
His Gln Pro Gln Thr Glu Glu Ala Lys Lys Glu Phe Thr Gln Arg Asp
          405          410          415
Val Gln Glu Lys Phe Asp Ile Thr Leu Thr Lys His Lys Gly Gln Tyr
          420          425          430
Ala Leu Arg Lys Leu Tyr His Glu Ser Pro Arg Val Ile Leu Val Leu
          435          440          445
Tyr Thr Ser Pro Thr Cys Gly Pro Cys Arg Thr Leu Lys Pro Ile Leu
          450          455          460
Asn Lys Val Val Asp Glu Tyr Asn His Asp Val His Phe Val Glu Ile
          465          470          475          480
Asp Ile Glu Glu Asp Gln Glu Ile Ala Glu Ala Ala Gly Ile Met Gly
          485          490          495
Thr Pro Cys Val Gln Phe Phe Lys Asn Lys Glu Met Leu Arg Arg Gly
          500          505          510
Asn Val Leu Ser Val Leu Lys Leu His Arg Leu Leu Cys Ser Gly Leu
          515          520          525
Ala Lys Asp Ser Glu Ser Val
          530          535

```

<210> 233
 <211> 117
 <212> PRT
 <213> *Helianthus annuus*

```

<400> 233
Ala Val Val Glu Ala Tyr Gly Glu Glu Gly Lys Asn Val Leu Gly Gly
1      5      10      15
Leu Lys Val Lys Asn Val Val Ser Gly Glu Val Ser Asp Leu Lys Val
20      25      30
Asn Gly Leu Phe Phe Ala Ile Gly His Glu Pro Ala Thr Lys Phe Leu
35      40      45
Asp Gly Gln Leu Glu Leu Asp Ser Asp Gly Tyr Val Thr Lys Pro
50      55      60
Gly Thr Thr Ile Ser Ser Val Lys Gly Val Phe Ala Ala Gly Asp Val
65      70      75      80
Gln Asp Lys Lys Tyr Arg Gln Ala Val Thr Ala Ala Gly Ser Gly Cys
85      90      95
Met Ala Ala Leu Asp Ala Glu His Tyr Leu Gln Glu Ile Gly Ser Gln
100      105      110
Glu Gly Lys Ser Asp
115

```

<210> 234
 <211> 300
 <212> PRT
 <213> *Arcaeoglobus fulgidus*

```

<400> 234
Met Tyr Asp Val Ala Ile Ile Gly Gly Gly Pro Ala Gly Leu Thr Ala
1      5      10      15
Ala Leu Tyr Ser Ala Arg Tyr Gly Leu Lys Thr Val Phe Phe Glu Thr
20      25      30
Val Asp Pro Val Ser Gln Leu Ser Leu Ala Ala Lys Ile Glu Asn Tyr
35      40      45
Pro Gly Phe Glu Gly Ser Gly Met Glu Leu Leu Glu Lys Met Lys Glu
50      55      60

```

Gln Ala Val Lys Ala Gly Ala Glu Trp Lys Leu Glu Lys Val Glu Arg
 65 70 75 80
 Val Glu Arg Asn Gly Glu Thr Phe Thr Val Ile Ala Glu Gly Gly Glu
 85 90 95
 Tyr Glu Ala Lys Ala Ile Ile Val Ala Thr Gly Gly Lys His Lys Glu
 100 105 110
 Ala Gly Ile Glu Gly Glu Ser Ala Phe Ile Gly Arg Gly Val Ser Tyr
 115 120 125
 Cys Ala Thr Cys Asp Gly Asn Phe Phe Arg Gly Lys Lys Val Ile Val
 130 135 140
 Tyr Gly Ser Gly Lys Glu Ala Ile Glu Asp Ala Ile Tyr Leu His Asp
 145 150 155 160
 Ile Gly Cys Glu Val Thr Ile Val Ser Arg Thr Pro Ser Phe Arg Ala
 165 170 175
 Glu Lys Ala Leu Val Glu Glu Val Glu Lys Arg Gly Ile Pro Val His
 180 185 190
 Tyr Ser Thr Thr Ile Arg Lys Ile Gly Ser Gly Lys Val Glu Lys
 195 200 205
 Val Val Ala Tyr Asn Arg Glu Lys Lys Glu Glu Phe Glu Ile Glu Ala
 210 215 220
 Asp Gly Ile Phe Val Ala Ile Gly Met Arg Pro Ala Thr Asp Val Val
 225 230 235 240
 Ala Glu Leu Gly Val Glu Arg Asp Ser Met Gly Tyr Ile Lys Val Asp
 245 250 255
 Lys Glu Gln Arg Thr Asn Val Glu Gly Val Phe Ala Ala Gly Asp Cys
 260 265 270
 Cys Asp Asn Pro Leu Lys Gln Val Val Thr Ala Cys Gly Asp Gly Ala
 275 280 285
 Val Ala Ala Tyr Ser Ala Tyr Lys Tyr Leu Thr Ser
 290 295 300

<210> 235

<211> 315

<212> PRT

<213> Bacillus halodurans

<400> 235

Met Gly Glu Glu Gln Lys Val Tyr Asp Val Val Ile Ala Gly Ala Gly
 1 5 10 15
 Pro Ala Gly Met Thr Ala Ala Val Tyr Thr Ser Arg Ala Asn Leu Ser
 20 25 30
 Thr Val Met Val Glu Arg Gly Val Pro Gly Gly Gln Met Ala Asn Thr
 35 40 45
 Glu Asp Val Glu Asn Tyr Pro Gly Phe Asp His Ile Leu Gly Pro Glu
 50 55 60
 Leu Ser Thr Lys Met Phe Glu His Ala Lys Lys Phe Gly Ala Glu Tyr
 65 70 75 80
 Ala Tyr Gly Asp Ile Lys Glu Ile Ile Asp Gln Gly Asp Leu Lys Leu
 85 90 95
 Val Lys Ala Gly Asn Lys Glu Tyr Lys Ala Arg Ala Val Ile Val Ala
 100 105 110
 Thr Gly Ala Glu Tyr Lys Lys Leu Gly Val Pro Gly Glu Lys Glu Leu
 115 120 125
 Ser Gly Arg Gly Val Ser Tyr Cys Ala Val Cys Asp Gly Ala Phe Phe
 130 135 140
 Lys Gly Lys Glu Leu Val Val Gly Gly Gly Asp Ser Ala Val Glu
 145 150 155 160
 Glu Ala Val Tyr Leu Thr Arg Phe Ala Ser Lys Val Thr Ile Ile His
 165 170 175
 Arg Arg Asp Gln Leu Arg Ala Gln Lys Ile Leu Gln Gln Arg Ala Phe
 180 185 190
 Asp Asn Asp Lys Ile Glu Phe Ile Trp Asp His Val Val Lys Gln Ile
 195 200 205
 Asn Gly Thr Asp Gly Lys Val Ser Ser Val Thr Ile Glu His Ala Lys
 210 215 220
 Thr Gly Glu Gln Gln Asp Phe Lys Thr Asp Gly Val Phe Ile Tyr Ile

225					230				235				240
Gly	Met	Leu	Pro	Leu	Asn	Glu	Ala	Val	Lys	Asn	Leu	Asn	Ile
				245					250				255
Asp	Glu	Gly	Tyr	Ile	Val	Thr	Asn	Glu	Glu	Met	Glu	Thr	Ser
			260					265					270
Gly	Ile	Phe	Ala	Ala	Gly	Asp	Val	Arg	Glu	Lys	Ser	Leu	Arg
		275					280					285	
Val	Thr	Ala	Thr	Gly	Asp	Gly	Ser	Leu	Ala	Ala	Gln	Asn	Val
	290					295					300		
Tyr	Ile	Glu	Glu	Leu	Ala	Glu	Lys	Val	Lys	Asn			
305					310					315			

<210> 236
 <211> 330
 <212> PRT
 <213> Bacillus halodurans

<400> 236													
Met	Ser	Arg	Lys	Glu	Glu	Leu	Tyr	Asp	Ile	Thr	Ile	Ile	Gly
1				5				10					15
Pro	Thr	Gly	Leu	Phe	Ala	Ala	Phe	Tyr	Gly	Gly	Met	Arg	Gln
			20					25					30
Val	Lys	Ile	Ile	Glu	Ser	Met	Pro	Gln	Leu	Gly	Gly	Gln	Leu
		35					40					45	
Leu	Tyr	Pro	Glu	Lys	Tyr	Ile	Tyr	Asp	Val	Ala	Gly	Phe	Pro
	50					55					60		Lys
Lys	Ala	Gln	Asp	Leu	Val	Asn	Asp	Leu	Lys	Arg	Gln	Ala	Glu
65				70						75			Gln
Asn	Pro	Thr	Ile	Ala	Leu	Glu	Gln	Ser	Val	Gln	Asn	Val	Thr
				85				90					Lys
Thr	Asp	Asp	Thr	Phe	Thr	Ile	Lys	Thr	Asp	Lys	Glu	Thr	His
			100					105					Tyr
Lys	Ala	Ile	Ile	Ile	Thr	Ala	Gly	Ala	Gly	Ala	Phe	Gln	Pro
	115						120					125	Arg
Leu	Glu	Val	Glu	Gly	Ala	Lys	Gln	Tyr	Glu	Gly	Lys	Asn	Leu
	130					135					140		Gln
Phe	Val	Asn	Asp	Leu	Asn	Ala	Tyr	Ala	Gly	Lys	Asn	Val	Leu
145					150					155			Ile
Gly	Gly	Gly	Asp	Ser	Ala	Val	Asp	Trp	Ala	Leu	Met	Leu	Glu
			165						170				Pro
Ala	Lys	Asn	Val	Thr	Leu	Ile	His	Arg	Arg	Asp	Lys	Phe	Arg
			180					185					Ala
Glu	His	Ser	Val	Glu	Leu	Leu	Gln	Lys	Ser	Ser	Val	Asn	Ile
	195						200					205	Leu
Pro	Phe	Ala	Ile	Ser	Glu	Leu	Ser	Gly	Asp	Gly	Glu	Lys	Ile
	210					215					220		His
Val	Thr	Ile	Gln	Glu	Val	Lys	Gly	Asp	Ala	Val	Glu	Thr	Leu
225					230					235			Asp
Asp	Glu	Val	Ile	Val	Asn	Phe	Gly	Phe	Val	Ser	Ser	Leu	Gly
				245					250				Pro
Lys	Gly	Trp	Gly	Leu	Glu	Ile	Glu	Lys	Asn	Ser	Ile	Val	Val
			260					265					Asn
Lys	Met	Glu	Thr	Asn	Ile	Pro	Gly	Ile	Tyr	Ala	Ala	Gly	Asp
		275					280					285	Ile
Thr	Tyr	Pro	Gly	Lys	Val	Lys	Leu	Ile	Ala	Thr	Gly	Phe	Gly
	290					295					300		Glu
Pro	Thr	Ala	Val	Asn	Asn	Ala	Lys	Ala	Phe	Ile	Asp	Pro	Thr
305					310					315			Ala
Val	Phe	Pro	Gly	His	Ser	Thr	Ser	Leu	Phe				Arg
				325					330				320

<210> 237
 <211> 213
 <212> PRT
 <213> Bacillus halodurans

<400> 237

```

Met Thr Asn Leu His Tyr Thr Val Lys Ser Leu Met Arg Phe Lys Asp
 1          5          10          15
Lys Thr Val Ile Ile Ser Gly Gly Gly Asn Ser Ala Ile Asp Trp Ala
 20          25          30
Asn Glu Leu Glu Pro Ile Ala Lys Val Tyr Leu Thr Tyr Arg Lys
 35          40          45
Glu Ala Leu Asn Gly His Glu Ala Gln Ile Ser Gln Leu Leu Ser Ser
 50          55          60
Ser Ala Thr Cys Leu Phe His Thr Thr Ile Ser Lys Leu Ile Ala Arg
 65          70          75          80
Asp Asn Lys Glu Val Ile Glu Gln Val Glu Leu Thr Asp His Gln Thr
 85          90          95
Gly Glu Val Thr Asn Leu Ala Val Asp Glu Val Ile Ile Asn His Gly
100          105          110
Tyr Glu Arg Asp Lys Ser Leu Leu Asp Gln Ser Glu Val Thr Leu Asp
115          120          125
Arg Ile Asp Asp Tyr Tyr Ile Ala Gly Thr Pro Thr Ser Ala Thr Ser
130          135          140
Val Gly Gly Ile Tyr Ala Ala Gly Asp Val Leu Lys His Glu Gly Lys
145          150          155          160
Leu His Leu Ile Ala Gly Ala Phe Gln Asp Ala Ala Asn Ala Val Asn
165          170          175
Gln Ala Lys Gln Trp Ile Glu Pro Glu Ala His Gln Ser Ala Met Val
180          185          190
Ser Ser His Asn His Val Phe Lys Glu Arg Asn Arg Glu Leu Ile Arg
195          200          205
Gln Met Leu Lys Asn
210

```

<210> 238

<211> 136

<212> PRT

<213> *Bacillus halodurans*

<400> 238

```

Met Asn Trp Glu Glu Leu Tyr Asp Val Thr Ile Ile Gly Gly Gly Pro
 1          5          10          15
Ala Gly Leu Phe Ser Ala Phe Tyr Ser Gly Leu Arg Glu Met Lys Thr
 20          25          30
Lys Val Ile Glu Tyr Gln Pro Met Leu Gly Gly Lys Val His Val Tyr
 35          40          45
Pro Glu Lys Met Ile Trp Asp Val Gly Gly Leu Thr Pro Ile Leu Gly
 50          55          60
Glu Lys Leu Ile Glu Gln Leu Val Thr Gln Ala Leu Thr Phe Asn Pro
 65          70          75          80
Thr Val Val Leu Asn Glu Lys Val Thr Ser Ile Ala Gln Glu Glu Ser
 85          90          95
Gly Trp Phe Val Ile Arg Thr Ala Ser Gly Arg Ala His Leu Thr Lys
100          105          110
Thr Val Ile Ile Ala Val Gly Gly Ile Leu Lys Pro Gln Lys Asn
115          120          125
Arg Ala Arg Arg Gly Arg Thr Ile
130          135

```

<210> 239

<211> 312

<212> PRT

<213> *Campylobacter jejuni*

<400> 239

```

Met Leu Asp Val Ala Ile Ile Gly Gly Gly Pro Ala Gly Leu Ser Ala
 1          5          10          15
Gly Leu Tyr Ala Thr Arg Gly Gly Leu Lys Asn Val Val Met Phe Glu
 20          25          30

```

Lys Gly Met Pro Gly Gly Gln Ile Thr Ser Ser Ser Glu Ile Glu Asn
 35 40 45
 Tyr Pro Gly Val Ala Gln Val Met Asp Gly Ile Ser Phe Met Ala Pro
 50 55 60
 Trp Ser Glu Gln Cys Met Arg Phe Gly Leu Lys His Glu Met Val Gly
 65 70 75 80
 Val Glu Gln Ile Leu Lys Asn Ser Asp Gly Ser Phe Thr Ile Lys Leu
 85 90 95
 Glu Gly Gly Lys Thr Glu Leu Ala Lys Ala Val Ile Val Cys Thr Gly
 100 105 110
 Ser Ala Pro Lys Lys Ala Gly Phe Lys Gly Glu Asp Glu Phe Phe Gly
 115 120 125
 Lys Gly Val Ser Thr Cys Ala Thr Cys Asp Gly Phe Phe Tyr Lys Asn
 130 135 140
 Lys Glu Val Ala Val Leu Gly Gly Gly Asp Thr Ala Leu Glu Glu Ala
 145 150 155 160
 Leu Tyr Leu Ala Asn Ile Cys Ser Lys Ile Tyr Leu Ile His Arg Arg
 165 170 175
 Asp Glu Phe Arg Ala Ala Pro Ser Thr Val Glu Lys Val Lys Lys Asn
 180 185 190
 Glu Lys Ile Glu Leu Ile Thr Ser Ala Ser Val Asp Glu Val Tyr Gly
 195 200 205
 Asp Lys Met Gly Val Ala Gly Val Lys Val Lys Leu Lys Asp Gly Ser
 210 215 220
 Ile Arg Asp Leu Asn Val Pro Gly Ile Phe Thr Phe Val Gly Leu Asn
 225 230 235 240
 Val Arg Asn Glu Ile Leu Lys Gln Asp Asp Ser Lys Phe Leu Cys Asn
 245 250 255
 Met Glu Glu Gly Gly Gln Val Ser Val Asp Leu Lys Met Gln Thr Ser
 260 265 270
 Val Ala Gly Leu Phe Ala Ala Gly Asp Leu Arg Lys Asp Ala Pro Lys
 275 280 285
 Gln Val Ile Cys Ala Ala Gly Asp Gly Ala Val Ala Ala Leu Ser Ala
 290 295 300
 Met Ala Tyr Ile Glu Ser Leu His
 305 310

<210> 240

<211> 348

<212> PRT

<213> *Caulobacter crescentus*

<400> 240

Met Ser Pro Leu Arg Arg Ile His Thr Ile Ser Pro Pro Met Ser Thr
 1 5 10 15
 Leu Ser Pro Arg Gln Thr Arg Cys Leu Ile Ile Gly Ser Gly Pro Ala
 20 25 30
 Gly Tyr Thr Ala Ala Ile Tyr Ala Ala Arg Ala Leu Leu Lys Pro Val
 35 40 45
 Leu Ile Ala Gly Ile Gln Pro Gly Gly Gln Leu Thr Ile Thr Thr Asp
 50 55 60
 Val Glu Asn Tyr Pro Gly Phe Ala Asp Val Ile Gln Gly Pro Trp Leu
 65 70 75 80
 Met Asp Gln Met Arg Ala Gln Ala Glu His Val Gly Thr Glu Phe Val
 85 90 95
 Ser Asp Ile Val Thr Ser Val Asp Leu Ser Lys Arg Pro Phe Thr Val
 100 105 110
 Lys Thr Asp Ser Gly Gln Asp Trp Ile Ala Glu Thr Ile Ile Ile Ala
 115 120 125
 Thr Gly Ala Gln Ala Lys Trp Leu Gly Leu Glu Ser Glu Ala Lys Phe
 130 135 140
 Gln Gly Phe Gly Val Ser Ala Cys Ala Thr Cys Asp Gly Phe Phe Tyr
 145 150 155 160
 Arg Asn Lys Asp Val Ile Val Val Gly Gly Gly Asn Thr Ala Val Glu
 165 170 175
 Glu Ala Leu Phe Leu Thr Ser Phe Ala Ser Lys Val Thr Leu Val His


```

      180      185      190
Arg Lys Asp Glu Leu Arg Ala Glu Lys Ile Leu Gln Glu Arg Leu Leu
      195      200      205
Ala His Pro Lys Ile Glu Val Ile Trp Asp Ser Val Ile Asp Glu Val
      210      215      220
Leu Gly Gln Thr Asp Pro Met Gly Val Thr Gly Ala Arg Leu Lys Asn
      225      230      235
Val Lys Thr Gly Glu Thr Gln Glu Val Ala Ala Asp Gly Val Phe Ile
      245      250      255
Ala Ile Gly His Ala Pro Ser Ser Glu Leu Phe Ala Gly Gln Leu Glu
      260      265      270
Thr Gly Ser Gly Gly Tyr Leu Lys Val Lys Pro Gly Thr Ala Ser Thr
      275      280      285
Ala Ile Glu Gly Val Tyr Ala Ala Gly Asp Val Thr Asp Asp Val Tyr
      290      295      300
Arg Gln Ala Val Thr Ala Ala Gly Met Gly Cys Met Ala Ala Leu Glu
      305      310      315
Ala Val Arg Phe Leu Ala Glu Glu Asp His Lys Ala Ala His His Pro
      325      330      335
Ile Ser His Ala Glu Ala Asn Lys Ile Gly Val Trp
      340      345

```

<210> 241

<211> 285

<212> PRT

<213> Clostridium acetobutylicum

<400> 241

```

Met Glu Arg Tyr Asp Ile Ala Ile Ile Gly Ser Gly Pro Ala Gly Leu
 1      5      10      15
Ala Ser Ala Ile Asn Ala Lys Thr Arg Asn Lys Ser Val Ile Val Phe
      20      25      30
Gly Ser Ser Asp Leu Ser Lys Lys Leu Thr Leu Ala Pro Val Ile Asn
      35      40      45
Asn Tyr Leu Gly Phe Tyr Gly Ile Arg Gly Ala Glu Leu Gln Glu Lys
      50      55      60
Phe Lys Glu His Ile Asp Asn Met Gly Ile Gln Ile Glu Asn Val Lys
      65      70      75      80
Val Asn Asn Ile Tyr Ala Met Gly Glu Tyr Phe Ser Ile Met Thr Ser
      85      90      95
Lys Asp Thr Tyr Glu Ala Ser Lys Val Ile Leu Ala Met Gly Met Glu
      100      105      110
His Thr Lys Pro Leu Lys Gly Glu Asp Lys Phe Leu Gly Arg Gly Val
      115      120      125
Gly Tyr Cys Ala Thr Cys Asp Ala Pro Leu Tyr Lys Gly Lys Ile Val
      130      135      140
Thr Ile Val Gly Tyr Asn Lys Glu Ala Glu Ser Glu Ala Asn Tyr Leu
      145      150      155      160
Ala Glu Leu Ala Ser Lys Val Tyr Tyr Val Pro Arg Tyr Lys Asp Glu
      165      170      175
Tyr Gln Leu Val Ser Ala Val Glu Ile Val Lys Asp Val Pro Val Glu
      180      185      190
Ile Val Gly Asp Lys Lys Val Glu Lys Leu Lys Leu Lys Ser Arg Glu
      195      200      205
Leu Glu Thr Asp Gly Val Phe Val Leu Lys Asp Ser Ala Pro Pro Glu
      210      215      220
Gln Leu Val Pro Gly Leu Tyr Val Glu Asp Gly His Ile Lys Val Asn
      225      230      235
Arg Lys Met Glu Thr Asn Ile Asp Gly Cys Tyr Ala Ala Gly Asp Cys
      245      250      255
Thr Gly Lys Pro Tyr Gln Tyr Met Lys Ala Val Gly Glu Gly Gln Val
      260      265      270
Ala Ala Leu Asn Ala Val Glu Lys Leu Tyr Thr Lys Ala
      275      280      285

```

<210> 242
 <211> 291
 <212> PRT
 <213> Clostridium acetobutylicum

<400> 242
 Met Asp Arg Tyr Asp Ile Ala Ile Ile Gly Ser Gly Pro Ala Gly Leu
 1 5 10 15
 Ser Ala Ala Ile Asn Ala Val Ile Arg Asn Lys Lys Val Ile Leu Phe
 20 25 30
 Gly Ser Asp Asn Leu Ser Asn Lys Leu Leu Lys Ala Pro Lys Ile Asn
 35 40 45
 Asn Tyr Leu Gly Ile Tyr Asp Val Ser Gly Lys Glu Leu Lys Glu Lys
 50 55 60
 Phe Leu Glu His Leu Lys Tyr Met Asn Ile Glu Ile Lys Asn Glu Lys
 65 70 75 80
 Val Asn Ser Val Tyr Ser Met Gly Asp Tyr Phe Ala Leu Ser Leu Asn
 85 90 95
 Gln Lys Met Tyr Glu Ala Thr Ser Ile Ile Ile Ala Ser Gly Val Glu
 100 105 110
 Phe Ser Lys Pro Leu Asn Gly Glu Asp Glu Leu Leu Gly Lys Gly Val
 115 120 125
 Gly Tyr Cys Ala Thr Cys Asp Ala Pro Leu Tyr Lys Gly Lys Thr Val
 130 135 140
 Ala Ile Val Gly Tyr Thr Lys Glu Ala Glu Glu Glu Ala Asn Tyr Val
 145 150 155 160
 Ser Glu Leu Ala Gly Lys Leu Tyr Tyr Ile Pro Met Tyr Lys Asp Lys
 165 170 175
 Val Ser Leu Lys Glu Val Ile Glu Val Val Glu Asp Lys Pro Ile Ser
 180 185 190
 Ile Leu Gly Lys Asp Lys Val Ser Gly Leu Gln Met Ser Lys Gly Glu
 195 200 205
 Ile Asn Thr Asp Ala Val Phe Ile Ile Lys Asp Ser Val Ser Pro Gly
 210 215 220
 Lys Leu Val Pro Gly Leu Leu Met Asn Gly Glu His Ile Ala Val Asp
 225 230 235 240
 Ile Asp Met Lys Thr Asn Ile Glu Gly Cys Phe Ala Ala Gly Asp Cys
 245 250 255
 Ala Gly Arg Pro Tyr Gln Tyr Ile Lys Ser Ala Gly Gln Gly Gln Ile
 260 265 270
 Ala Ala Leu Ser Ala Val Ser Tyr Ile Asp Lys Ile Lys Leu Asn Lys
 275 280 285
 Lys Ile Ile
 290

<210> 243
 <211> 314
 <212> PRT
 <213> Clostridium sticklandii

<400> 243
 Met Ser Lys Ile Tyr Asp Leu Val Ile Ile Gly Ala Gly Pro Ala Gly
 1 5 10 15
 Leu Ser Ala Gly Leu Tyr Gly Ala Arg Gly Lys Met Ser Thr Leu Ile
 20 25 30
 Ile Glu Lys Asp Lys Thr Gly Gly Gln Ile Val Thr Thr Glu Glu Val
 35 40 45
 Ala Asn Tyr Pro Gly Ser Ile His Asp Ala Ser Gly Pro Ser Leu Ile
 50 55 60
 Ala Arg Met Ala Glu Gln Ala Asp Glu Phe Gly Thr Glu Arg Ile Lys
 65 70 75 80
 Asp Ser Ile Val Asp Phe Asp Phe Thr Gly Lys Ile Lys Ile Leu Lys
 85 90 95
 Gly Thr Lys Ala Glu Tyr Gln Ala Lys Ala Val Ile Val Ala Thr Gly
 100 105 110
 Ala Ser Pro Lys Lys Leu Asp Cys Pro Gly Glu Lys Glu Leu Thr Gly

Lys	Gly	115	Val	Ser	Tyr	Cys	Ala	120	Thr	Cys	Asp	Ala	Asp	125	Phe	Phe	Gln	Asp
130	Met	Glu	Val	Phe	Val	Val	Gly	135	Gly	Gly	Asp	Ser	Ala	140	Val	Glu	Glu	Ala
145	Met	Tyr	Leu	Thr	Lys	Phe	Ala	150	Ser	Lys	Val	Thr	Ile	155	Val	His	Arg	160
165	Asp	Ser	Leu	Arg	Ala	Ala	Lys	170	Ser	Ile	Gln	Asp	Lys	175	Ala	Phe	Ala	Asn
180	Pro	Lys	Ile	Asp	Phe	Lys	Trp	185	Asp	Ser	Val	Ile	Lys	190	Glu	Ile	Lys	Gly
195	Asp	Gly	Ile	Val	Glu	Ser	Val	200	Phe	Glu	Asn	Thr	Lys	205	Thr	Gly	Glu	
210	Leu	Ser	Glu	His	Phe	Ala	Asp	215	Glu	Glu	Phe	Gly	Thr	220	Phe	Gly	Ile	Phe
225	Val	Phe	Thr	Gly	Tyr	Ile	Pro	230	Gln	Thr	Asp	Ile	Phe	235	Lys	Asp	Lys	Val
245	Asp	Met	Asn	Gln	Ser	Gly	Tyr	250	Phe	Val	Thr	Asn	Gln	255	Asn	Met	Glu	Thr
260	Asn	Ile	Pro	Gly	Val	Phe	Ala	265	Ala	Gly	Asp	Cys	Arg	270	Glu	Lys	Val	Leu
275	Arg	Gln	Val	Val	Thr	Ala	Thr	280	Ala	Asp	Gly	Ala	Ile	285	Ala	Ala	Ile	Met
290	Ala	Glu	Lys	Tyr	Ile	Glu	His	295	Glu	Gly	Leu			300				
305								310										

<210> 244

<211> 325

<212> PRT

<213> *Deinococcus radiodurans*

<400> 244

Met	Thr	Ala	Pro	Thr	Ala	His	Asp	Tyr	Asp	Val	Val	Ile	Ile	Gly	Gly
1				5					10					15	
Gly	Pro	Ala	Gly	Leu	Thr	Ala	Ala	Ile	Tyr	Thr	Gly	Arg	Ala	Gln	Leu
			20					25					30		
Ser	Thr	Leu	Ile	Leu	Glu	Lys	Gly	Met	Pro	Gly	Gly	Gln	Ile	Ala	Trp
		35					40					45			
Ser	Glu	Glu	Val	Glu	Asn	Phe	Pro	Gly	Phe	Pro	Glu	Pro	Ile	Ala	Gly
		50				55					60				
Met	Glu	Leu	Ala	Gln	Arg	Met	His	Gln	Gln	Ala	Glu	Lys	Phe	Gly	Ala
65					70					75					80
Lys	Val	Glu	Met	Asp	Glu	Val	Gln	Gly	Val	Gln	His	Asp	Ala	Thr	Ser
				85				90					95		
His	Pro	Tyr	Pro	Phe	Thr	Val	Arg	Gly	Tyr	Asn	Gly	Glu	Tyr	Arg	Ala
			100					105					110		
Lys	Ala	Val	Ile	Leu	Ala	Thr	Gly	Ala	Asp	Pro	Arg	Lys	Leu	Gly	Ile
		115					120					125			
Pro	Gly	Glu	Asp	Asn	Phe	Trp	Gly	Lys	Gly	Val	Ser	Thr	Cys	Ala	Thr
		130				135					140				
Cys	Asp	Gly	Phe	Phe	Tyr	Lys	Gly	Lys	Lys	Val	Val	Ile	Gly	Gly	
145					150					155				160	
Gly	Asp	Ala	Ala	Val	Glu	Glu	Gly	Met	Phe	Leu	Thr	Lys	Phe	Ala	Asp
				165				170					175		
Glu	Val	Thr	Val	Ile	His	Arg	Arg	Asp	Thr	Leu	Arg	Ala	Asn	Lys	Val
			180					185					190		
Ala	Gln	Ala	Arg	Ala	Phe	Ala	Asn	Pro	Lys	Met	Lys	Phe	Ile	Trp	Asp
		195					200					205			
Thr	Ala	Val	Glu	Glu	Ile	Gln	Gly	Ala	Asp	Ser	Val	Ser	Gly	Val	Lys
		210				215					220				
Leu	Arg	Asn	Leu	Lys	Thr	Gly	Glu	Val	Ser	Glu	Leu	Ala	Thr	Asp	Gly
225					230					235				240	
Val	Phe	Ile	Phe	Ile	Gly	His	Val	Pro	Asn	Thr	Ala	Phe	Val	Lys	Asp
				245				250					255		
Thr	Val	Ser	Leu	Arg	Asp	Asp	Gly	Tyr	Val	Asp	Val	Arg	Asp	Glu	Ile
			260					265					270		

Tyr Thr Asn Ile Pro Met Leu Phe Ala Ala Gly Asp Val Ser Asp Tyr
 275 280 285
 Ile Tyr Arg Gln Leu Ala Thr Ser Val Gly Ala Gly Thr Arg Ala Ala
 290 295 300
 Met Met Thr Glu Arg Gln Leu Ala Ala Leu Glu Val Glu Gly Glu Glu
 305 310 315 320
 Val Thr Ala Ala Asp
 325

<210> 245
 <211> 61
 <212> PRT
 <213> Enterococcus faecalis

<220>
 <221> VARIANT
 <222> 33, 45, 46
 <223> Xaa = Any Amino Acid

<400> 245
 Met Met Asp Thr Leu Ile Ile Glu Lys Asp Lys Ile Gly Gly Gln Val
 1 5 10 15
 Thr Thr Thr Ser Glu Ile Val Asn Tyr Pro Ala Ile Arg His Thr Thr
 20 25 30
 Xaa Pro Glu Leu Met Gly Glu Met Arg Ile Gln Ala Xaa Xaa Phe Gly
 35 40 45
 Val Ala Phe Thr Lys Asp Glu Ile Ile Asp Val Asp Phe
 50 55 60

<210> 246
 <211> 205
 <212> PRT
 <213> Halobacterium sp

<400> 246
 Met Thr Glu Asp Ser His Asp Leu Val Ile Ala Gly Ser Gly Ile Ala
 1 5 10 15
 Gly Leu Ser Ala Ala Val Tyr Ala Ala Arg Ala Asp Leu Glu Pro Leu
 20 25 30
 Val Leu Glu Gly Asp Glu Pro Gly Gly Gln Leu Thr Leu Thr Thr Asp
 35 40 45
 Val Glu Asn Tyr Leu Gly Phe Pro Asp Gly Val Gly Gly Met Asp Leu
 50 55 60
 Val Gln Arg Gly Lys Glu Gln Ala Glu Gln Phe Gly Ala Gln Phe Glu
 65 70 75 80
 His Gly Arg Ile Glu Ala Ala Asp Leu Asp Gly Gln Pro Leu Glu Leu
 85 90 95
 Ser Leu Ser Thr Gly Asp Thr Leu Tyr Thr Arg Ser Leu Ile Val Ala
 100 105 110
 Thr Gly Ala Ser Ala Arg Trp Val Gly Ala Glu Asn Glu Asp Glu Leu
 115 120 125
 Met Gly Ala Gly Leu Ser Thr Cys Ala Thr Cys Asp Gly Ala Phe His
 130 135 140
 Arg Gly Asp Asp Val Leu Val Val Gly Gly Gly Asp Ser Ala Met Glu
 145 150 155 160
 Glu Ala Leu Phe Leu Ala Lys Phe Ala Asp Ser Val Thr Val Val His
 165 170 175
 Arg Arg Glu Glu Leu Arg Ala Ser Glu Ile Met Ala Asp Arg Ala Arg
 180 185 190
 Asp His Asp Asp Val Gln Phe Arg Trp Asn Thr Glu Leu
 195 200 205

<210> 247
 <211> 362

<212> PRT

<213> Halobacterium sp

<400> 247

```

Met Thr Glu Ala Thr Ala Asp Arg Thr Ala Leu Thr Asp Gly Gly Arg
 1          5          10          15
Asp Val Val Glu His Arg Gln Leu Val Ile Val Gly Ser Gly Ile Ala
 20          25          30
Ala Leu Ser Ala Ala Thr Tyr Ala Ala Arg Ser Asn Asn Asp Pro Leu
 35          40          45
Leu Phe Glu Gly Asp Glu Pro Gly Gly Gln Leu Thr Leu Thr Ser Glu
 50          55          60
Val Glu Asn Tyr Pro Gly Phe Pro Glu Gly Ile Ala Gly Ala Glu Leu
 65          70          75          80
Ile Gln Glu Met Lys Thr Gln Ala Thr Arg Phe Gly Ala Glu Val Glu
 85          90          95
His Gly Ile Val Glu Ser Val Asp Asp Ser Gly Arg Pro Phe Arg Leu
 100         105         110
Thr Leu Thr Asn Gly Asp Val Tyr Thr Ala Asp Ala Val Ile Val Ala
 115         120         125
Ser Gly Ala Ser Ala Arg Thr Leu Gly Ile Pro Gly Glu Asp Glu Leu
 130         135         140
Met Gly Gln Gly Val Ser Thr Cys Ala Thr Cys Asp Gly Ala Phe Phe
 145         150         155         160
Arg Gly Glu Asp Met Ile Val Val Gly Gly Gly Asp Ala Ala Ala Glu
 165         170         175
Glu Ala Ser Phe Leu Thr Lys Phe Ala Asp Thr Val Tyr Leu Val His
 180         185         190
Arg Arg Asp Glu Leu Arg Ala Glu Asp Tyr Trp Ala Asp Arg Ile Arg
 195         200         205
Glu His Val Ala Asp Gly Asp Ile Glu Val Leu Trp Asn Thr Glu Ala
 210         215         220
Val Glu Val His Gly Ser Pro Glu Glu Gly Val Thr Gly Ala Ser Leu
 225         230         235
Val Arg His Pro Glu Gly His Pro Thr Ala Lys Leu Asp Ala Asp Glu
 245         250         255
Thr Glu Gln Leu Glu Leu Asp Ile Gly Ala Phe Phe Ile Ala Ile Gly
 260         265         270
His Thr Pro Asn Thr Ser Phe Leu Ala Asp Thr Gly Val Val Cys Asp
 275         280         285
Asp Ala Gly Tyr Val Gln Thr Val Gly Gly Ala Gly Gly Gly Gln Thr
 290         295         300
Lys Thr Asp Val Thr Gly Val Phe Gly Ala Gly Asp Val Val Asp Tyr
 305         310         315
His Tyr Gln Gln Ala Val Thr Ala Ala Gly Met Gly Ser Lys Ala Ala
 325         330         335
Ile Asp Ala Asp Glu Tyr Leu Glu Ser Val Ala Asp Gly Val Thr Gly
 340         345         350
Glu Thr Ala Asp Ala Thr Pro Ala Asp Asp
 355         360

```

<210> 248

<211> 294

<212> PRT

<213> Halobacterium

<400> 248

```

Met Pro Thr Gln Asp Gly Glu Arg Arg Asp Val Val Ile Val Gly Gly
 1          5          10          15
Gly Pro Ala Gly Cys Ala Ala Gly Val Phe Thr Ala Arg Tyr Gly Leu
 20          25          30
Asp Thr Val Val Phe Asp Arg Gly Asn Ala Ala Leu Pro Arg Cys Ala
 35          40          45
Phe Val Glu Asn Tyr Pro Gly Phe Pro Gly Gly Ile Asp Val Pro Thr
 50          55          60
Leu Arg Gly Leu Phe His Asp His Ala Glu Thr Ala Gly Cys Asp Leu

```

65	Ile	Ala	Asp	Thr	Val	70	Glu	Ser	Val	Asp	Arg	75	Pro	Ser	Asp	Asp	Asp	80	Thr
					85						90						95		
	Gly	Phe	Val	Val	Glu	Thr	Gln	Asp	Gly	Arg	Arg	Val	Tyr	Thr	Thr	Asp	Thr		
				100					105						110				
	Val	Leu	Ala	Ala	Ala	Trp	Tyr	Asp	Gly	Ser	Tyr	Leu	Arg	Pro	Val	Val			
			115					120						125					
	Gly	Asp	Ser	Ala	Phe	Glu	Thr	His	Asp	His	His	Gly	Glu	Ser	Arg	Glu			
			130					135					140						
	Arg	Phe	Asp	Asp	Ala	Tyr	Ala	Asp	Ala	Asp	Gly	Arg	Thr	Pro	Val	Asp			
			145				150					155				160			
	Gly	Leu	Tyr	Val	Ala	Ser	Pro	Gly	Gly	Gln	Arg	Ser	Ala	Gln	Ala	Val			
				165						170					175				
	Ile	Ala	Ala	Gly	Asn	Gly	Ala	His	Val	Ala	Arg	Cys	Leu	Leu	Ala	Asp			
				180					185					190					
	Arg	Lys	Arg	Ala	Arg	Gly	Tyr	Pro	Glu	Gly	Val	Ala	Pro	His	Tyr	Asp			
			195					200					205						
	Trp	Lys	Arg	Arg	Glu	Ser	Asp	Leu	Ser	Gly	Glu	Trp	Ala	Asp	Arg	Asp			
			210				215					220							
	Arg	Trp	Arg	Glu	Trp	Phe	Ala	Ala	Glu	Ala	Gly	Asp	Asp	His	Asp	Leu			
			225			230					235				240				
	Asp	Asp	Asp	Glu	Phe	Ala	Ala	Leu	Arg	Ala	Ala	His	Leu	Asp	Arg	Thr			
				245					250					255					
	Phe	Asp	Ala	Thr	Leu	Ser	Ala	Asp	Ala	Ile	Glu	Glu	Arg	Ala	Glu	Ala			
				260					265					270					
	Gly	Ala	His	Arg	Leu	Leu	Asp	His	Ile	Asp	Asp	Asp	His	Ile	Glu	Ser			
			275					280					285						
	Tyr	Arg	Glu	Gln	Arg	Asp													
			290																

<210> 249

<211> 324

<212> PRT

<213> Helicobacter pylori

<400> 249

Met	Asn	Gln	Glu	Ile	Leu	Asp	Val	Leu	Ile	Val	Gly	Ala	Gly	Pro	Gly
1				5					10					15	
Gly	Ile	Ala	Thr	Ala	Val	Glu	Cys	Glu	Ile	Ala	Gly	Val	Lys	Lys	Val
			20					25					30		
Leu	Leu	Cys	Glu	Lys	Thr	Glu	Ser	His	Ser	Gly	Met	Leu	Glu	Lys	Phe
		35				40					45				
Tyr	Lys	Ala	Gly	Lys	Arg	Ile	Asp	Lys	Asp	Tyr	Lys	Lys	Gln	Val	Val
	50				55				60						
Glu	Leu	Lys	Gly	His	Ile	Pro	Phe	Lys	Asp	Ser	Phe	Lys	Glu	Glu	Thr
65				70					75					80	
Leu	Glu	Asn	Phe	Thr	Asn	Leu	Leu	Lys	Glu	His	His	Ile	Thr	Pro	Ser
			85					90					95		
Tyr	Lys	Thr	Asp	Ile	Glu	Ser	Val	Lys	Lys	Glu	Gly	Glu	Tyr	Phe	Lys
			100					105					110		
Ile	Thr	Thr	Thr	Ser	Asn	Thr	Thr	Tyr	His	Ala	Lys	Phe	Val	Val	Val
		115					120					125			
Ala	Ile	Gly	Lys	Met	Gly	Gln	Pro	Asn	Arg	Pro	Thr	Ala	Tyr	Lys	Ile
	130					135				140					
Pro	Val	Ala	Leu	Ser	Lys	Gln	Val	Val	Phe	Ser	Ile	Asn	Asp	Cys	Lys
145					150					155				160	
Glu	Asn	Glu	Lys	Thr	Leu	Val	Ile	Gly	Gly	Asn	Ser	Ala	Val	Glu	
			165					170					175		
Tyr	Ala	Ile	Ala	Leu	Cys	Lys	Thr	Thr	Pro	Thr	Thr	Leu	Asn	Tyr	Arg
			180					185					190		
Lys	Lys	Glu	Phe	Ser	Arg	Ile	Asn	Glu	Asp	Asn	Ala	Lys	Asn	Leu	Gln
		195					200				205				
Glu	Val	Leu	Asn	Asn	Asn	Thr	Leu	Lys	Ser	Lys	Leu	Gly	Val	Asp	Ile
	210					215				220					
Glu	Ser	Leu	Glu	Glu	Asp	Asn	Thr	Gln	Ile	Lys	Val	Asn	Phe	Thr	Asp
225					230					235					240

Asn Thr Ser Glu Ser Phe Asp Arg Leu Leu Tyr Ala Ile Gly Gly Ser
 245 250 255
 Thr Pro Leu Glu Phe Phe Lys Arg Cys Ser Leu Glu Leu Asp Pro Ser
 260 265 270
 Thr Asn Ile Pro Val Val Lys Glu Asn Leu Glu Ser Asn Asn Ile Pro
 275 280 285
 Asn Leu Phe Ile Val Gly Asp Ile Leu Phe Lys Ser Gly Ala Ser Ile
 290 295 300
 Ala Thr Ala Leu Asn His Gly Tyr Asp Val Ala Ile Glu Ile Ala Lys
 305 310 315 320
 Arg Leu His Ser

<210> 250
 <211> 128
 <212> PRT
 <213> Klebsiella oxytoca

<400> 250
 Met Gly Thr Ala Lys His Ser Lys Leu Leu Ile Leu Gly Ser Gly Pro
 1 5 10 15
 Ala Gly Tyr Thr Ala Ala Val Tyr Ala Arg Ala Asn Leu Gln Pro
 20 25 30
 Val Leu Ile Thr Gly Met Glu Lys Gly Gly Gln Leu Thr Thr Thr Thr
 35 40 45
 Glu Val Glu Asn Trp Pro Gly Asp Pro Asn Asp Leu Thr Gly Pro Leu
 50 55 60
 Leu Met Glu Arg Met His Glu His Ala Thr Lys Phe Glu Thr Glu Ile
 65 70 75 80
 Ile Phe Asp His Ile Asn Ser Val Asp Leu Gln Asn Arg Pro Phe Arg
 85 90 95
 Leu Val Gly Asp Ser Gly Glu Tyr Thr Cys Asp Ala Pro Asp Tyr Arg
 100 105 110
 Tyr Arg Arg Ile Ser Ala Leu Ser Gly Ser Ala Ile Gly Arg Arg Val
 115 120 125

<210> 251
 <211> 79
 <212> PRT
 <213> Lactococcus lactis

<400> 251
 Met Gln Glu Leu Asp Leu Ile Ile Val Gly Ala Gly Pro Val Gly Leu
 1 5 10 15
 Tyr Ala Ala Phe Tyr Ala Gly Met Arg Gly Leu Ser Val Ala Ile Ile
 20 25 30
 Glu Ser Ala Gln Val Pro Gly Gly Gln Pro Gln Asn Leu Tyr Pro Glu
 35 40 45
 Lys Leu Ile Tyr Asp Ile Ala Gly Leu Pro Ala Val Thr Gly Ala Asp
 50 55 60
 Leu Thr Lys Asn Leu Leu Glu Gln Leu Ala Gln Ile Ser His Arg
 65 70 75

<210> 252
 <211> 321
 <212> PRT
 <213> Lactococcus lactis

<400> 252
 Met Gln Glu Leu Asp Leu Ile Ile Val Gly Ala Gly Pro Val Gly Leu
 1 5 10 15
 Tyr Ala Ala Phe Tyr Ala Gly Met Arg Gly Leu Ser Val Ala Ile Ile
 20 25 30
 Glu Ser Ala Gln Val Pro Gly Gly Gln Pro Gln Asn Leu Tyr Pro Glu

35 40 45
 Lys Leu Ile Tyr Asp Ile Ala Gly Leu Pro Ala Val Thr Gly Ala Asp
 50 55 60
 Leu Thr Lys Asn Leu Leu Glu Gln Leu Ala Gln Ile Ser His Arg Leu
 65 70 75 80
 Phe Leu Gly Glu Ser Val Gln Lys Ile Glu Lys Glu Glu Gly Ile Phe
 85 90 95
 Ser Val Thr Thr Asp Lys Ser Thr Arg Arg Ala Lys Gly Val Leu Leu
 100 105 110
 Thr Thr Gly Ala Gly Leu Leu Lys Pro Arg Lys Leu Gly Ile Asp Asn
 115 120 125
 Glu Glu Thr Leu Ala Asn Glu Gly Lys Ile Ser Tyr Phe Ile Thr Ser
 130 135 140
 Leu Lys Glu Phe Glu Gly Lys Asn Val Ala Val Phe Gly Gly Gly Asp
 145 150 155 160
 Ser Ala Leu Asp Trp Ser Leu Met Leu Glu Lys Val Ala Lys Asn Val
 165 170 175
 His Leu Val His Arg Arg Thr Ala Phe Arg Gly His Glu Ile Thr Val
 180 185 190
 Asp Arg Val Met Asn Ser Asn Val Gln Val His Thr Pro Tyr Thr Phe
 195 200 205
 Ser Asn Leu Ile Glu Asn Glu Leu Glu Leu Lys Lys Ile Lys Ser Glu
 210 215 220
 Glu Ser Leu Asn Phe Ser Ile Asp Lys Ile Leu Val Asn Tyr Gly Phe
 225 230 235 240
 Leu Thr Asn Gln Val Thr Leu Ala Glu Asn Leu Glu Val Ser Arg Asn
 245 250 255
 Gly Arg Val Lys Ala Asp Ser Met Met Gln Ser Asn Ile Glu Gly Leu
 260 265 270
 Tyr Val Ala Gly Asp Ala Ser Asp Tyr Pro Gly Lys Met Pro Leu Met
 275 280 285
 Ser Val Gly Phe Gly Glu Ala Val His Ala Ile Asn Ala Met Thr Lys
 290 295 300
 Lys Leu Glu Phe Asp His Pro Leu Arg Gly Gly His Ser Ser Ser Ile
 305 310 315 320
 Phe

<210> 253
 <211> 308
 <212> PRT
 <213> Lactococcus lactis

<400> 253
 Met Thr Glu Lys Lys Tyr Asp Val Val Ile Ile Gly Ser Gly Pro Ala
 1 5 10 15
 Gly Met Thr Ala Ala Met Tyr Thr Ala Arg Ser Glu Met Lys Thr Leu
 20 25 30
 Leu Leu Glu Arg Gly Val Pro Gly Gly Gln Met Asn Asn Thr Ala Glu
 35 40 45
 Ile Glu Asn Tyr Pro Gly Tyr Glu Thr Ile Met Gly Pro Glu Leu Ser
 50 55 60
 Met Lys Met Ala Glu Pro Leu Glu Gly Leu Gly Val Glu Asn Ala Tyr
 65 70 75 80
 Gly Phe Val Thr Ala Ile Glu Asp His Gly Asp Tyr Lys Lys Ile Ile
 85 90 95
 Thr Glu Asp Asp Glu Phe Val Thr Lys Ser Ile Ile Ala Thr Gly
 100 105 110
 Ala Asn His Arg Lys Leu Glu Ile Pro Gly Glu Glu Glu Tyr Gly Ala
 115 120 125
 Arg Gly Val Ser Tyr Cys Ala Val Cys Asp Gly Ala Phe Phe Arg Asn
 130 135 140
 Gln Glu Ile Leu Val Ile Gly Gly Gly Asp Ser Ala Val Glu Glu Ala
 145 150 155 160
 Leu Tyr Leu Thr Arg Phe Gly Gln Ser Val Thr Ile Met His Arg Arg
 165 170 175

Asp Lys Leu Arg Ala Gln Glu Ile Ile Gln Gln Arg Ala Phe Lys Glu
 180 185 190
 Glu Lys Ile Asn Phe Ile Trp Asp Ser Val Pro Met Glu Ile Lys Gly
 195 200 205
 Asp Asp Lys Lys Val Gln Ser Val Val Tyr Lys Asn Val Lys Thr Gly
 210 215 220
 Glu Val Thr Glu Lys Ala Phe Gly Gly Ile Phe Ile Tyr Val Gly Leu
 225 230 235 240
 Asp Pro Val Ala Glu Phe Ala Gly Asn Leu Gly Ile Thr Asp Glu Ala
 245 250 255
 Gly Trp Ile Ile Thr Asp Asp His Met Arg Thr Ser Leu Pro Gly Ile
 260 265 270
 Phe Ala Val Gly Asp Val Arg Gln Lys Asp Phe Arg Gln Ile Thr Thr
 275 280 285
 Ala Ile Gly Asp Gly Ala Gln Ala Ala Gln Glu Ala Tyr Lys Phe Val
 290 295 300
 Ala Glu Leu Asp
 305

<210> 254
 <211> 44
 <212> PRT
 <213> *Lactococcus lactis*

<400> 254
 Met Gln Glu Leu Asp Leu Ile Ile Val Gly Ala Gly Pro Val Gly Leu
 1 5 10 15
 Tyr Ala Ala Phe Tyr Ala Gly Met Arg Gly Leu Ser Val Ala Ile Ile
 20 25 30
 Glu Ser Ala Gln Val Pro Gly Gly Gln Pro Gln Asn
 35 40

<210> 255
 <211> 339
 <212> PRT
 <213> *Listeria monocytogenes*

<400> 255
 Glu Phe Tyr Ser Tyr Lys Lys Glu Ile Asn Arg Tyr Leu Ala Glu Glu
 1 5 10 15
 Asp Ser Ala Ser Ala Cys Asp Ile Leu Arg Lys Val Ile Asp Glu Lys
 20 25 30
 Pro Asn Phe Trp Pro Ala Tyr Asn Gln Leu Ala Ser Leu Tyr Phe Glu
 35 40 45
 Gln Leu Lys Glu Glu Glu Gly Val Arg Val Leu Ser Asp Leu Leu Ser
 50 55 60
 Arg Asn Pro Gly Asn Leu Leu Gly Ile Cys Asp Leu Phe Ile Tyr His
 65 70 75 80
 Phe Tyr Lys Gly Asn Arg Lys Glu Ala Asp Glu Leu Tyr Leu Glu Leu
 85 90 95
 Arg Asp Val Leu Pro Val Leu Ala His His Lys Glu Lys Leu Gly Leu
 100 105 110
 Ile His Ala Met Met Gly Glu Tyr Glu Glu Ala Asp Asp Leu Leu Glu
 115 120 125
 Gln Val Ala Asp Leu Glu Val Thr Glu Arg Ser Lys Tyr Tyr Tyr Phe
 130 135 140
 Arg Ala Lys Ser Ser Tyr Tyr Leu Gly Asp Val Glu Gly Ala Lys Met
 145 150 155 160
 Phe Trp His Ser Phe Leu Glu Cys Asp Leu Tyr Glu Asp Val Arg Phe
 165 170 175
 Pro Trp Glu Gln Glu Pro Asp Leu Thr Asn Asp Thr Arg Leu Val Leu
 180 185 190
 Glu Met Leu Gln Glu Glu Asp Asp Leu Thr His Met Leu Gly Val Tyr
 195 200 205
 Ala Leu Thr Ile Ser Gly Asn Arg Pro Glu Leu Val Leu Phe His Pro

210	215	220
Leu Leu Asp Met Ser Asp Trp Ser Tyr Met Glu His Leu Met Phe Thr		
225	230	235
Asn Phe Asp Tyr Phe Pro Asp Gly Ala Ile Glu Gln Asn Gly Tyr Leu		
	245	250
Ile Ala Lys Ala Met Ile Ile Leu Lys Glu Asn Gly Ile Leu Leu Asn		
	260	265
Glu Glu Tyr Met Ala Leu Tyr Lys Gln Met Phe Ser Leu Val Leu Ile		
	275	280
Asp Ala Gly Lys Asp Leu Ile Leu Gly Arg Tyr Thr Ile Glu Thr Val		
	290	295
Ala Ser Ala Ile Ala Lys Leu Phe Leu Pro His Leu Lys Leu Gln Leu		
305	310	315
Val Glu Glu Phe Glu Cys Ser Lys Cys Ala Arg Asp Ile Glu Arg Val		
	325	330
		335
Leu Ser Arg		

<210> 256

<211> 303

<212> PRT

<213> Methanothermobacter thermautotrophicus

<400> 256

Met Met Thr Asp Tyr Asp Met Ile Val Ile Gly Ala Gly Pro Ala Gly	
1	5
Leu Thr Ala Gly Ile Tyr Gly Gly Arg Gln Gly Ser Ser Val Leu Met	
	20
Leu Asp Lys Gly Pro Ala Gly Gly Leu Gly Leu Glu Val Pro Met Met	
	35
Glu Asn Tyr Pro Gly Phe Glu Met Ile Ala Gly Met Ser Leu Val Thr	
	50
Lys Met Lys Lys Gln Ala Thr Ala Val Ala Glu Arg Glu Met Glu	
65	70
Glu Val Lys Glu Ile Glu Lys Gly Asp Val Phe Thr Val Lys Thr Ser	
	85
Arg Asp Thr Tyr Thr Ala Ser Ala Ile Phe Ala Thr Gly Ser Lys	
	100
His Arg Gln Leu Gly Val Pro Gly Glu Asn Asp Leu Leu Gly Arg Gly	
	115
Val Cys Tyr Cys Ala Thr Cys Asp Gly Pro Leu Tyr Lys Gly Arg Lys	
	130
Val Leu Met Val Gly Gly Asn Ser Ala Ala Gln Glu Ala Val Phe	
145	150
Leu Lys Asn Ile Gly Cys Asp Val Ser Ile Val His Arg Arg Asp Glu	
	165
Leu Arg Ala Asp Lys Tyr Leu Gln Asp Lys Leu Arg Glu Met Glu Ile	
	180
Pro Val Ile Trp Asn Ser Val Val Lys Glu Ile Gly Gly Asp Glu Arg	
	195
Val Glu Glu Val Ile Ile His Asn Arg Val Thr Gly Arg Asp Glu Thr	
	210
Leu Lys Val Asp Gly Val Phe Ile Ala Ile Gly Glu Glu Pro Leu Asn	
225	230
Gln Leu Ala Val Asp Leu Gly Val Glu Val Asp Lys Gly Gly Tyr Ile	
	245
Ile Thr Asp Lys Phe Gln Arg Thr Asn Val Pro Leu Val Tyr Ala Ala	
	260
Gly Asp Ile Thr Gly Gly Leu Asn Gln Trp Val Thr Ala Cys Ala Glu	
	275
Gly Ala Ile Ala Ala Thr Tyr Ala Tyr Arg Glu Ile Gln Ser Tyr	
	290
	295
	300

<210> 257

<211> 179

<212> PRT

<213> *Bacillus subtilis*

<400> 257

```

Met Val Ile Ser Gly Gly Gly Asp Thr Ala Val Asp Trp Ala Asn Glu
 1          5          10          15
Leu Glu Pro Ile Ala Ala Ser Val Thr Val Val His Arg Arg Glu Glu
          20          25          30
Phe Gly Gly Met Glu Ser Ser Val Thr Lys Met Lys Gln Ser Ser Val
          35          40          45
Arg Val Leu Thr Pro Tyr Arg Leu Glu Gln Leu Asn Gly Asp Glu Glu
          50          55          60
Gly Ile Lys Ser Val Thr Val Cys His Thr Glu Ser Gly Gln Arg Lys
          65          70          75          80
Asp Ile Glu Ile Asp Glu Leu Ile Ile Asn His Gly Phe Lys Ile Asp
          85          90          95
Leu Gly Pro Met Met Glu Trp Gly Leu Glu Ile Glu Glu Gly Arg Val
          100          105          110
Lys Ala Asp Arg His Met Arg Thr Asn Leu Pro Gly Val Phe Val Ala
          115          120          125
Gly Asp Ala Ala Phe Tyr Glu Ser Lys Leu Arg Leu Ile Ala Gly Gly
          130          135          140
Phe Thr Glu Gly Pro Thr Ala Val Asn Ser Ala Lys Ala Tyr Leu Asp
          145          150          155          160
Pro Lys Ala Glu Asn Met Ala Met Tyr Ser Thr His His Lys Lys Leu
          165          170          175
Val His Lys

```

<210> 258

<211> 307

<212> PRT

<213> *Mycoplasma pulmonis*

<400> 258

```

Met Ser Gln Asn Lys Ile Tyr Asp Val Ala Ile Ile Gly Ala Gly Pro
 1          5          10          15
Gly Ala Leu Thr Ala Ala Ile Tyr Thr Ser Arg Gly Asn Leu Asp Thr
          20          25          30
Val Phe Ile Asp Asn Ala Ala Pro Gly Gly Lys Leu Ile Tyr Ala Ser
          35          40          45
Lys Ile Glu Asn Trp Pro Gly Asp Thr Ile Val Lys Gly Thr Asp Leu
          50          55          60
Ala Ile Arg Phe Phe Glu His Ala Gln Ala Phe Gly Ala Lys Tyr Glu
          65          70          75          80
Tyr Gly Lys Val Val Asp Leu Ile Asn Ile Lys Asp Asp Leu Lys Glu
          85          90          95
Leu Val Leu Glu Asp Gly Lys Lys Ile Gln Ala Lys Ser Val Ile Ile
          100          105          110
Ala Ser Gly Met Val Ser Arg Lys Pro Arg Glu Ile Leu Asn Tyr Asp
          115          120          125
Glu Phe Glu Asn Arg Gly Val Ser Tyr Cys Val Ile Cys Asp Gly Pro
          130          135          140
Met Tyr Gly His Asn Pro Ala Ile Ile Ile Gly Gly Gly Asn Ser Ala
          145          150          155          160
Val Glu Glu Gly Thr Phe Leu Ser Ser Ile Ala Ser Lys Val Tyr Val
          165          170          175
Ile Val Arg Asp Ser Asp Phe Ile Ala Glu Lys Ala Leu Val Asn Asp
          180          185          190
Leu Lys Ser Arg Lys Asn Ile Glu Val Leu Phe Asn Ala Ser Val Lys
          195          200          205
Glu Leu His Gly Lys Asp Ala Leu Glu Tyr Ala Ile Val Asn His Asn
          210          215          220
Gly Lys Glu Val Lys Leu Glu Val Ala Ser Leu Phe Pro Tyr Ile Gly
          225          230          235          240
Phe Leu Pro Ser Ala Glu Tyr Ala Lys Asn Ala Gly Val Leu Glu Pro

```

```

                245                250                255
Asn Gly Phe Ile Lys Thr Asp Glu Phe Met Glu Thr Lys Val Pro Gly
                260                265                270
Ile Tyr Ala Ile Gly Asp Ile Arg Ile Lys Asp Ile Arg Gln Ile Leu
                275                280                285
Thr Ala Thr Ser Asp Gly Thr Ile Ala Gly Lys Ile Leu Thr Asn Arg
                290                295                300
Ile Lys Lys
305

```

<210> 259

<211> 316

<212> PRT

<213> Neisseria meningitidis

<400> 259

```

Met Ser Gln His Arg Lys Leu Ile Ile Leu Gly Ser Gly Pro Ala Gly
1          5          10          15
Tyr Thr Ala Ala Val Tyr Ala Ala Arg Ala Asn Leu Asn Pro Val Ile
          20          25          30
Ile Thr Gly Ile Ala Gln Gly Gly Gln Leu Met Thr Thr Thr Glu Val
          35          40          45
Asp Asn Trp Pro Ala Asp Ala Asp Gly Val Gln Gly Thr Glu Leu Met
          50          55          60
Ala Arg Phe Leu Ala His Ala Glu Arg Phe Gly Thr Glu Ile Ile Phe
          65          70          75          80
Asp Gln Ile Asn Ala Val Asp Leu Gln Lys Arg Pro Phe Thr Leu Lys
          85          90          95
Gly Asp Met Gly Glu Tyr Thr Cys Asp Ala Leu Ile Val Ala Thr Gly
          100          105          110
Ala Ser Ala Lys Tyr Leu Gly Leu Pro Ser Glu Glu Ala Phe Ala Gly
          115          120          125
Lys Gly Val Ser Ala Cys Ala Thr Cys Asp Gly Phe Phe Tyr Lys Asn
          130          135          140
Gln Asp Val Ala Val Val Gly Gly Gly Asn Thr Ala Val Glu Glu Ala
          145          150          155          160
Leu Tyr Leu Ala Asn Ile Ala Lys Thr Val Thr Leu Ile His Arg Arg
          165          170          175
Ser Glu Phe Arg Ala Glu Lys Ile Met Ile Asp Lys Leu Met Lys Arg
          180          185          190
Val Glu Glu Gly Lys Ile Ile Leu Lys Leu Glu Ser Asn Leu Gln Glu
          195          200          205
Val Leu Gly Asp Asp Arg Gly Val Asn Gly Ala Leu Leu Lys Asn Asn
          210          215          220
Asp Gly Ser Glu Gln Gln Ile Ala Val Ser Gly Ile Phe Ile Ala Ile
          225          230          235          240
Gly His Lys Pro Asn Thr Asp Ile Phe Lys Gly Gln Leu Glu Met Asp
          245          250          255
Glu Ala Gly Tyr Leu Lys Thr Lys Gly Gly Thr Ala Asp Asn Val Gly
          260          265          270
Ala Thr Asn Ile Glu Gly Val Trp Ala Ala Gly Asp Val Lys Asp His
          275          280          285
Thr Tyr Arg Gln Ala Ile Thr Ser Ala Ala Ser Gly Cys Gln Ala Ala
          290          295          300
Leu Asp Ala Glu Arg Trp Leu Gly Ser Gln Asn Ile
305          310          315

```

<210> 260

<211> 316

<212> PRT

<213> Neisseria meningitidis

<400> 260

```

Met Ser Gln His Arg Lys Leu Ile Ile Leu Gly Ser Gly Pro Ala Gly
1          5          10          15

```

Tyr Thr Ala Ala Val Tyr Ala Ala Arg Ala Asn Leu Asn Pro Val Ile
 20 25 30
 Ile Thr Gly Ile Ala Gln Gly Gly Gln Leu Met Thr Thr Thr Glu Val
 35 40 45
 Asp Asn Trp Pro Ala Asp Ala Asp Gly Val Gln Gly Pro Glu Leu Met
 50 55 60
 Ala Arg Phe Leu Ala His Ala Glu Arg Phe Gly Thr Glu Ile Ile Phe
 65 70 75 80
 Asp Gln Ile Asn Ala Val Asp Leu Gln Lys Arg Pro Phe Thr Leu Lys
 85 90 95
 Gly Asp Met Gly Glu Tyr Thr Cys Asp Ala Leu Ile Val Ala Thr Gly
 100 105 110
 Ala Ser Ala Lys Tyr Leu Gly Leu Pro Ser Glu Glu Ala Phe Ala Gly
 115 120 125
 Lys Gly Val Ser Ala Cys Ala Thr Cys Asp Gly Phe Phe Tyr Lys Asn
 130 135 140
 Gln Asp Val Ala Val Val Gly Gly Gly Asn Thr Ala Val Glu Glu Ala
 145 150 155 160
 Leu Tyr Leu Ala Asn Ile Ala Lys Thr Val Thr Leu Ile His Arg Arg
 165 170 175
 Ser Glu Phe Arg Ala Glu Lys Ile Met Ile Asp Lys Leu Met Lys Arg
 180 185 190
 Val Glu Glu Gly Lys Ile Ile Leu Lys Leu Glu Ser Asn Leu Gln Glu
 195 200 205
 Val Leu Gly Asp Asp Arg Gly Val Asn Gly Ala Leu Leu Lys Asn Asn
 210 215 220
 Asp Gly Ser Glu Gln Gln Ile Ala Val Ser Gly Ile Phe Ile Ala Ile
 225 230 235 240
 Gly His Lys Pro Asn Thr Asp Ile Phe Lys Gly Gln Leu Glu Met Asp
 245 250 255
 Glu Ala Gly Tyr Leu Lys Thr Lys Gly Gly Thr Ala Asp Asn Val Gly
 260 265 270
 Ala Thr Asn Ile Glu Gly Val Trp Ala Ala Gly Asp Val Lys Asp His
 275 280 285
 Thr Tyr Arg Gln Ala Ile Thr Ser Ala Ala Ser Gly Cys Gln Ala Ala
 290 295 300
 Leu Asp Ala Glu Arg Trp Leu Gly Ser Gln Asn Ile
 305 310 315

<210> 261

<211> 316

<212> PRT

<213> *Pseudomonas aeruginosa*

<400> 261

Met Ser Glu Val Lys His Ser Arg Leu Ile Ile Leu Gly Ser Gly Pro
 1 5 10 15
 Ala Gly Tyr Thr Ala Ala Val Tyr Ala Ala Arg Ala Asn Leu Lys Pro
 20 25 30
 Val Val Ile Thr Gly Ile Gln Pro Gly Gly Gln Leu Thr Thr Thr Thr
 35 40 45
 Glu Val Asp Asn Trp Pro Gly Asp Val Glu Gly Leu Thr Gly Pro Ala
 50 55 60
 Leu Met Thr Arg Met Gln Gln His Ala Glu Arg Phe Asp Thr Glu Ile
 65 70 75 80
 Val Tyr Asp His Ile His Thr Ala Glu Leu Gln Gln Arg Pro Phe Thr
 85 90 95
 Leu Lys Gly Asp Ser Gly Thr Tyr Thr Cys Asp Ala Leu Ile Ile Ala
 100 105 110
 Thr Gly Ala Ser Ala Gln Tyr Leu Gly Met Ser Ser Glu Glu Ala Phe
 115 120 125
 Met Gly Lys Gly Val Ser Ala Cys Ala Thr Cys Asp Gly Phe Phe Tyr
 130 135 140
 Arg Asn Gln Val Val Cys Val Val Gly Gly Gly Asn Thr Ala Val Glu
 145 150 155 160
 Glu Ala Leu Tyr Leu Ala Asn Ile Ala Lys Glu Val His Leu Ile His

Arg	Arg	Asp	Lys	165	Leu	Arg	Ser	Glu	Lys	170	Ile	Leu	Gln	Asp	Lys	175	Leu	Phe
Asp	Lys	Ala	180	Glu	Asn	Gly	Asn	Val	185	His	Leu	His	Trp	Asn	190	Thr	Thr	Leu
Asp	Glu	Val	195	Leu	Gly	Asp	Ala	Ser	200	Gly	Val	Thr	Gly	Val	205	Arg	Leu	Lys
Ser	Thr	Ile	210	Asp	Gly	Ser	Thr	Ser	215	Glu	Leu	Ser	Leu	Ala	Gly	Val	Phe	
Ile	Ala	Ile	225	Gly	His	Lys	Pro	Asn	230	Thr	Asp	Leu	Phe	Gln	Gly	Gln	Leu	
Glu	Met	Arg	245	Asp	Gly	Tyr	Leu	Arg	250	Ile	His	Gly	Gly	Ser	Glu	Gly	Asn	
Ala	Thr	Gln	260	Thr	Ser	Ile	Glu	Gly	265	Val	Phe	Ala	Ala	Gly	270	Asp	Val	Ala
Asp	His	Val	275	Tyr	Arg	Gln	Ala	Ile	280	Thr	Ser	Ala	Gly	Ala	Gly	Cys	Met	
Ala	Ala	Leu	290	Asp	Ala	Glu	Lys	Tyr	295	Leu	Asp	Asp	His					
305					310						315							

<210> 262

<211> 316

<212> PRT

<213> Pseudomonas aeruginosa

<400> 262

Met	Pro	Asp	Thr	Leu	Arg	His	Ala	Arg	Val	Ile	Ile	Leu	Gly	Ser	Gly			
1				5					10					15				
Pro	Ala	Gly	Tyr	Ser	Ala	Ala	Val	Tyr	Ala	Ala	Arg	Ala	Asn	Leu	Lys			
			20					25					30					
Pro	Leu	Leu	Ile	Thr	Gly	Met	Gln	Ala	Gly	Gly	Gln	Leu	Thr	Thr	Thr			
		35					40					45						
Thr	Glu	Val	Asp	Asn	Trp	Pro	Gly	Asp	Pro	His	Gly	Leu	Thr	Gly	Pro			
	50					55					60							
Ala	Leu	Met	Gln	Arg	Met	Gln	Glu	His	Ala	Glu	Arg	Phe	Glu	Thr	Glu			
65					70					75					80			
Ile	Val	Phe	Asp	His	Ile	His	Ala	Val	Asp	Leu	Ala	Gly	Lys	Pro	Phe			
				85					90					95				
Thr	Leu	Arg	Gly	Asp	Asn	Gly	Thr	Tyr	Thr	Cys	Asp	Ala	Leu	Ile	Val			
			100					105					110					
Ala	Thr	Gly	Ala	Ser	Ala	Arg	Tyr	Leu	Gly	Leu	Pro	Ser	Glu	Gln	Ala			
		115					120					125						
Phe	Met	Gly	Lys	Gly	Val	Ser	Ala	Cys	Ala	Thr	Cys	Asp	Gly	Phe	Phe			
	130					135					140							
Tyr	Arg	Asn	Arg	Glu	Val	Ala	Val	Ile	Gly	Gly	Gly	Asn	Thr	Ala	Val			
145				150					155					160				
Glu	Glu	Ala	Leu	Tyr	Leu	Ala	Asn	Ile	Ala	Ser	Arg	Val	Thr	Leu	Val			
				165					170					175				
His	Arg	Arg	Glu	Thr	Phe	Arg	Ala	Glu	Lys	Ile	Leu	Gln	Asp	Lys	Leu			
			180					185					190					
Gln	Ala	Arg	Val	Ala	Glu	Gly	Lys	Ile	Val	Leu	Lys	Leu	Asn	Ala	Glu			
		195					200					205						
Val	Asp	Glu	Val	Leu	Gly	Asp	Thr	Met	Gly	Val	Thr	Gly	Val	Arg	Leu			
	210					215					220							
Lys	Thr	Arg	Asp	Gly	Gly	Ser	Glu	Glu	Ile	Ala	Val	Asp	Gly	Met	Phe			
225				230						235				240				
Val	Ala	Ile	Gly	His	Thr	Pro	Asn	Thr	Ser	Leu	Phe	Glu	Gly	Gln	Leu			
				245					250					255				
Ala	Leu	Lys	Asp	Gly	Tyr	Leu	Val	Val	Asn	Gly	Gly	Arg	Glu	Gly	Asn			
			260					265					270					
Ala	Thr	Ala	Thr	Asn	Val	Pro	Gly	Val	Phe	Ala	Ala	Gly	Asp	Val	Ala			
		275					280					285						
Asp	His	Val	Tyr	Arg	Gln	Ala	Ile	Thr	Ser	Ala	Gly	Ala	Gly	Cys	Met			
	290					295					300							
Ala	Ala	Leu	Asp	Val	Glu	Arg	Tyr	Leu	Asp	Ser	Leu							
305					310					315								

<210> 263
 <211> 345
 <212> PRT
 <213> *Pyrococcus abyssi*

<400> 263
 Met Leu Leu Asn Ile His Gln Glu Ser Tyr Val Glu Val Val Lys Met
 1 5 10 15
 Phe Ser Leu Gly Gly Leu Gly Lys Ser Arg Val Asp Glu Ser Lys Val
 20 25 30
 Trp Asp Val Ile Ile Ile Gly Ala Gly Pro Ala Gly Tyr Thr Ala Ala
 35 40 45
 Ile Tyr Ala Ala Arg Phe Gly Leu Asp Thr Ile Ile Ile Thr Lys Asp
 50 55 60
 Leu Gly Gly Asn Met Ala Ile Thr Asp Leu Ile Glu Asn Tyr Pro Gly
 65 70 75 80
 Phe Pro Glu Gly Ile Ser Gly Ser Glu Leu Ala Lys Arg Met Tyr Glu
 85 90 95
 His Val Lys Lys Tyr Gly Val Asp Val Ile Phe Asp Glu Val Val Arg
 100 105 110
 Ile Asp Pro Ala Glu Cys Ala Tyr Tyr Glu Gly Pro Cys Gln Phe Glu
 115 120 125
 Val Lys Thr Ala Asn Gly Lys Glu Tyr Lys Gly Lys Thr Ile Ile Ile
 130 135 140
 Ala Val Gly Ala Glu Pro Arg Lys Leu His Val Pro Gly Glu Lys Glu
 145 150 155 160
 Phe Thr Gly Arg Gly Val Ser Tyr Cys Ala Thr Cys Asp Gly Pro Leu
 165 170 175
 Phe Val Gly Lys Glu Val Ile Val Val Gly Gly Gly Asn Thr Ala Leu
 180 185 190
 Gln Glu Ala Leu Tyr Leu His Ser Ile Gly Val Lys Val Thr Leu Val
 195 200 205
 His Arg Arg Asp Lys Phe Arg Ala Asp Lys Ile Leu Gln Asp Arg Leu
 210 215 220
 Lys Gln Ala Gly Ile Pro Thr Ile Leu Asn Thr Val Val Thr Glu Ile
 225 230 235 240
 Arg Gly Thr Asn Lys Val Glu Ser Val Val Leu Lys Asn Val Lys Thr
 245 250 255
 Gly Glu Thr Phe Glu Lys Lys Val Asp Gly Val Phe Ile Phe Ile Gly
 260 265 270
 Tyr Glu Pro Lys Thr Asp Phe Val Lys His Leu Gly Ile Thr Asp Glu
 275 280 285
 Tyr Gly Tyr Ile Lys Val Asp Met Tyr Met Arg Thr Lys Val Pro Gly
 290 295 300
 Ile Phe Ala Ala Gly Asp Ile Thr Asn Val Phe Lys Gln Ile Ala Val
 305 310 315 320
 Ala Val Gly Gln Gly Ala Ile Ala Ala Asn Ser Ala Lys Glu Phe Ile
 325 330 335
 Glu Ser Trp Asn Gly Lys Ser Ile Glu
 340 345

<210> 264
 <211> 334
 <212> PRT
 <213> *Rickettsia prowazekii*

<400> 264
 Met Tyr Asn Thr Asp Ile Val Ile Ile Gly Ser Gly Pro Val Gly Leu
 1 5 10 15
 Phe Ala Val Phe Gln Ala Gly Met Leu Gly Met Lys Cys His Val Ile
 20 25 30
 Asp Ala Gln Glu Val Ile Gly Gly Gln Cys Ile Thr Leu Tyr Pro Glu
 35 40 45
 Lys His Ile Tyr Asp Ile Pro Ala Tyr Pro Lys Ile Ala Ala Lys Glu
 50 55 60

```

Leu Ile Lys Gln Leu Glu Ser Gln Ala Ala Pro Phe Asn Pro Val Tyr
65      70      75      80
His Leu Asn Gln Gln Ala Thr Glu Leu Asn Lys His Asp Asp Phe Phe
      85      90      95
Glu Ile Lys Thr Ser Lys Asn Thr Leu Ile Lys Ser Lys Val Ile Ile
      100      105      110
Ile Ala Ala Gly Ala Gly Ala Phe Gly Pro Asn Lys Pro Pro Ile Ala
      115      120      125
Asn Ile Glu Ala Phe Glu Gly Lys Ser Ile Phe Tyr Phe Ile Asn Asp
      130      135      140
Lys Ser Lys Phe Leu Gly Lys Asn Ile Val Val Ala Gly Gly Gly Asp
      145      150      155      160
Ser Ala Val Asp Trp Ala Ile Thr Leu Ser Glu Ile Ala Asn Lys Ile
      165      170      175
Tyr Leu Val His Arg Arg Asp Lys Phe Thr Ala Ala Thr Glu Ser Val
      180      185      190
Arg Gln Leu Arg His Ile Ala Glu Thr Gly Lys Ile Glu Leu Val Thr
      195      200      205
Gly Tyr Gln Leu Asn Asn Leu Asp Gly His Asn Ser Glu Leu Arg Ser
      210      215      220
Val Ile Val Lys Asp Leu Gln Asn Asn Ile Arg Lys Leu Asp Ala Asn
      225      230      235      240
Ile Leu Leu Pro Phe Gly Leu Lys Gln Asp Leu Gly Pro Leu Ala
      245      250      255
Asn Trp Gly Phe Asn Val Arg Leu Gln His Ile Glu Val Asp Asn Tyr
      260      265      270
Tyr Tyr Gln Thr Asn Ile Lys Gly Ile Tyr Ala Ile Gly Asp Val Ala
      275      280      285
His Tyr Val Gly Lys Leu Lys Leu Ile Ile Thr Gly Phe Ala Glu Ala
      290      295      300
Ala Cys Ser Leu His His Ala Tyr Ser Arg Val Phe Asp Gly Lys Ala
      305      310      315      320
Leu His Phe Glu Tyr Ser Thr Asn Lys Tyr Glu Gln Lys Gln
      325      330

```

<210> 265

<211> 311

<212> PRT

<213> Staphylococcus aureus

<400> 265

```

Met Thr Glu Ile Asp Phe Asp Ile Ala Ile Ile Gly Ala Gly Pro Ala
1      5      10      15
Gly Met Thr Ala Ala Val Tyr Ala Ser Arg Ala Asn Leu Lys Thr Val
      20      25      30
Met Ile Glu Arg Gly Ile Pro Gly Gly Gln Met Ala Asn Thr Glu Glu
      35      40      45
Val Glu Asn Phe Pro Gly Phe Glu Met Ile Thr Gly Pro Asp Leu Ser
      50      55      60
Thr Lys Met Phe Glu His Ala Lys Lys Phe Gly Ala Val Tyr Gln Tyr
      65      70      75      80
Gly Asp Ile Lys Ser Val Glu Asp Lys Gly Glu Tyr Lys Val Ile Asn
      85      90      95
Phe Gly Asn Lys Glu Leu Thr Ala Lys Ala Val Ile Ile Ala Thr Gly
      100      105      110
Ala Gly Tyr Lys Lys Ile Gly Val Pro Gly Glu Gln Glu Leu Gly Gly
      115      120      125
Arg Gly Val Ser Tyr Cys Ala Val Cys Asp Gly Ala Phe Phe Lys Asn
      130      135      140
Lys Arg Leu Phe Val Ile Gly Gly Gly Asp Ser Ala Val Glu Glu Gly
      145      150      155      160
Thr Phe Leu Thr Lys Phe Ala Asp Lys Val Thr Ile Val His Arg Arg
      165      170      175
Asp Glu Leu Arg Ala Gln Arg Ile Leu Gln Asp Arg Ala Phe Lys Asn
      180      185      190
Asp Lys Ile Asp Phe Ile Trp Ser His Thr Leu Lys Ser Ile Asn Glu

```


[illegible]

```
<210> 266
<211> 326
<212> PRT
<213> Streptomyces coelicolor
```

[illegible]

<210> 267

<211> 558

<212> PRT

<213> Streptomyces coelicolor

<400> 267

```

Met Ala Gln Ala Asp Gly Glu Thr Arg Thr Val Ile Met Thr Val Asp
1      5      10      15
Asp Asp Pro Gly Val Ser Arg Ala Val Ala Arg Asp Leu Arg Arg Arg
20      25      30
Tyr Gly Ala Thr Tyr Arg Ile Val Arg Ala Glu Ser Gly Glu Ser Ala
35      40      45
Leu Asp Ala Leu Arg Glu Leu Lys Leu Arg Gly Asp Leu Val Ala Val
50      55      60
Ile Leu Ala Asp Tyr Arg Met Pro Gln Met Asn Gly Ile Glu Phe Leu
65      70      75      80
Glu Gln Ala Leu Asp Val Tyr Pro Gly Ala Arg Arg Val Leu Leu Thr
85      90      95
Ala Tyr Ala Asp Thr Asn Ala Ala Ile Asp Ala Ile Asn Val Val Asp
100     105     110
Leu Asp His Tyr Leu Leu Lys Pro Trp Asp Pro Pro Glu Glu Lys Leu
115     120     125
Tyr Pro Val Leu Asp Asp Leu Leu Gln Ala Trp Arg Ala Gly Asp His
130     135     140
Arg Pro Val Pro Ser Thr Lys Val Val Gly His Arg Trp Ser Ala Arg
145     150     155     160
Ser Ser Glu Val Arg Glu Phe Leu Ala Arg Asn Gln Val Pro Tyr Arg
165     170     175
Trp Tyr Ser Ser Asp Glu Pro Glu Gly Arg Arg Leu Leu Ser Ala Ala
180     185     190
Gly Gln Asp Gly Gln Arg Leu Pro Val Val Ile Thr Pro Asp Gly Thr
195     200     205
Pro Leu Val Glu Pro Glu Ala Pro Glu Leu Ala Ala Arg Val Gly Leu
210     215     220
Ala Thr Thr Pro Thr Ser Asp Phe Tyr Asp Leu Val Val Ile Gly Gly
225     230     235     240
Gly Pro Ala Gly Leu Gly Ala Ala Val Tyr Gly Ala Ser Glu Gly Leu
245     250     255
Arg Thr Val Leu Val Glu Arg Ser Ala Thr Gly Gly Gln Ala Gly Gln
260     265     270
Ser Ser Arg Ile Glu Asn Tyr Leu Gly Phe Pro Asp Gly Val Ser Gly
275     280     285
Gly Gln Leu Thr Glu Arg Ala Arg Arg Gln Ala Ala Arg Phe Gly Ala
290     295     300
Glu Ile Leu Thr Ala Arg Glu Val Thr Gly Leu Glu Ala Asn Gly Ala
305     310     315     320
Ala Arg Val Val Arg Phe Ser Asp Gly Ser Ala Ile Ala Ala His Ser
325     330     335
Val Ile Leu Ala Thr Gly Val Ser Tyr Arg Gln Leu Thr Ala Pro Gly
340     345     350
Thr Glu Asp Leu Ala Gly Cys Gly Val Phe Tyr Gly Ser Ala Leu Thr
355     360     365
Glu Ala Ala Ser Cys Gln Gly His Asp Val Tyr Ile Val Gly Gly Ala
370     375     380
Asn Ser Ala Gly Gln Ala Ala Met Tyr Leu Ala Arg Gly Ala Lys Ser
385     390     395     400
Val Thr Leu Leu Val Arg Gly Gly Ser Leu Glu Ala Ser Met Ser Tyr
405     410     415
Tyr Leu Ile Gln Gln Ile Glu Glu Thr Pro Asn Ile Arg Val Arg Cys
420     425     430
Gly Thr Leu Val Glu Gly Ala His Gly Asp Gly His Leu Glu Arg Leu
435     440     445
Thr Leu Arg Asp Ala Ala Ser Gly Ala Thr Glu Leu Val Asp Ala Gln
450     455     460
Trp Leu Phe Val Phe Ile Gly Ala Ala Pro Leu Thr Asp Trp Leu Asp
465     470     475     480
Gly Thr Val Leu Arg Asp Glu Arg Gly Phe Ile Leu Ala Gly Pro Asp

```

				485					490				495				
Leu	Thr	Pro	Asp	Gly	Arg	Pro	Pro	Ala	Gly	Trp	Glu	Leu	Asp	Arg	Pro		
			500					505					510				
Pro	Tyr	His	Leu	Glu	Thr	Ser	Val	Pro	Gly	Val	Phe	Val	Ala	Gly	Asp		
		515					520					525					
Ala	Arg	Ala	Glu	Ser	Ala	Lys	Arg	Val	Ala	Ser	Ala	Val	Gly	Glu	Gly		
	530					535					540						
Ala	Met	Ala	Val	Met	Leu	Val	His	Arg	Tyr	Leu	Glu	Gln	Ser				
545					550					555							

<210> 268

<211> 303

<212> PRT

<213> Streptococcus pneumoniae

<400> 268

Met	Tyr	Asp	Thr	Ile	Ile	Ile	Gly	Ala	Gly	Pro	Ala	Gly	Met	Thr	Ala		
1				5					10				15				
Ala	Leu	Tyr	Ala	Ala	Arg	Ser	Asn	Leu	Lys	Val	Ala	Leu	Ile	Glu	Gly		
			20				25					30					
Gly	Leu	Pro	Gly	Gly	Gln	Met	Asn	Asn	Thr	Ser	Asp	Ile	Glu	Asn	Tyr		
		35				40						45					
Pro	Gly	Tyr	Ala	Asn	Ile	Ser	Gly	Pro	Glu	Leu	Ala	Glu	Lys	Met	Phe		
	50					55					60						
Glu	Pro	Leu	Glu	Asn	Leu	Gly	Val	Glu	His	Ile	Tyr	Gly	Tyr	Val	Glu		
65				70						75					80		
Asn	Val	Glu	Asp	His	Gly	Asp	Phe	Lys	Lys	Val	Met	Thr	Asp	Asp	Gln		
				85				90					95				
Thr	Tyr	Glu	Thr	Arg	Thr	Val	Ile	Val	Ala	Thr	Gly	Ser	Lys	His	Arg		
			100					105					110				
Pro	Leu	Gly	Val	Pro	Gly	Glu	Glu	Glu	Leu	Asn	Ser	Arg	Gly	Val	Ser		
		115					120					125					
Tyr	Cys	Ala	Val	Cys	Asp	Gly	Ala	Phe	Phe	Arg	Asp	Gln	Asp	Leu	Leu		
	130					135					140						
Val	Val	Gly	Gly	Gly	Asp	Ser	Ala	Val	Glu	Glu	Ala	Leu	Phe	Leu	Thr		
145					150					155					160		
Arg	Phe	Ala	Lys	Thr	Val	Thr	Ile	Val	His	Arg	Arg	Asp	Gln	Leu	Arg		
				165				170					175				
Ala	Gln	Lys	Val	Leu	Gln	Asp	Arg	Ala	Phe	Ala	Asn	Glu	Lys	Ile	Ser		
			180					185					190				
Phe	Ile	Trp	Asp	Ser	Val	Val	Arg	Glu	Ile	Lys	Gly	Glu	Asn	Arg	Val		
		195					200					205					
Glu	Ser	Val	Val	Phe	Glu	Asn	Val	Lys	Thr	Gly	Gln	Val	Thr	Glu	Gln		
		210				215					220						
Ala	Phe	Gly	Gly	Val	Phe	Ile	Tyr	Val	Gly	Leu	Asp	Pro	Leu	Ser	Asp		
					230					235					240		
Phe	Val	Lys	Glu	Leu	Asn	Ile	Gln	Asp	Gln	Ala	Gly	Trp	Ile	Val	Thr		
				245					250					255			
Asp	Asn	His	Met	Lys	Thr	Ala	Val	Asp	Gly	Ile	Phe	Ala	Val	Gly	Asp		
			260					265					270				
Val	Arg	Leu	Lys	Asp	Leu	Arg	Gln	Val	Thr	Thr	Ala	Val	Gly	Asp	Gly		
		275					280					285					
Ala	Ile	Ala	Gly	Gln	Glu	Ala	Tyr	Lys	Phe	Ile	Thr	Glu	His	Ser			
		290				295					300						

<210> 269

<211> 330

<212> PRT

<213> Streptococcus pyogenes

<400> 269

Met	Lys	Asp	Lys	Ala	Tyr	Asp	Ile	Thr	Ile	Ile	Gly	Gly	Gly	Pro	Ile		
1				5					10					15			
Gly	Leu	Phe	Ala	Ala	Phe	Tyr	Ala	Gly	Leu	Arg	Gly	Val	Thr	Val	Lys		
			20					25					30				

```

Ile Ile Glu Ser Leu Ser Glu Leu Gly Gly Gln Pro Ala Ile Leu Tyr
  35      40      45
Pro Glu Lys Met Ile Tyr Asp Ile Pro Ala Tyr Pro Ser Leu Thr Gly
  50      55      60
Val Glu Leu Thr Glu Asn Leu Ile Lys Gln Leu Ser Arg Phe Glu Asp
  65      70      75      80
Arg Thr Thr Ile Cys Leu Lys Glu Glu Val Leu Thr Phe Asp Lys Val
      85      90      95
Lys Gly Gly Phe Ser Ile Arg Thr Asn Lys Ala Glu His Phe Ser Lys
  100      105      110
Ala Ile Ile Ile Ala Cys Gly Asn Gly Ala Phe Ala Pro Arg Thr Leu
  115      120      125
Gly Leu Glu Ser Glu Glu Asn Phe Ala Asp His Asn Leu Phe Tyr Asn
  130      135      140
Val His Gln Leu Asp Gln Phe Ala Gly Gln Lys Val Val Ile Cys Gly
  145      150      155      160
Gly Gly Asp Ser Ala Val Asp Trp Ala Leu Ala Leu Glu Asp Ile Ala
      165      170      175
Glu Ser Val Thr Val Val His Arg Arg Asp Ala Phe Arg Ala His Glu
  180      185      190
His Ser Val Glu Leu Leu Lys Ala Ser Thr Val Asn Leu Leu Thr Pro
  195      200      205
Tyr Val Pro Lys Ala Leu Lys Gly Ile Gly Asn Leu Ala Glu Lys Leu
  210      215      220
Val Ile Gln Lys Val Lys Glu Asp Glu Val Leu Glu Leu Glu Leu Asp
  225      230      235      240
Ser Leu Ile Val Ser Phe Gly Phe Ser Thr Ser Asn Lys Asn Leu Lys
      245      250      255
Asn Trp Asn Leu Asp Tyr Lys Arg Ser Ser Ile Thr Val Ser Pro Leu
  260      265      270
Phe Gln Thr Ser Gln Glu Gly Ile Phe Ala Ile Gly Asp Ala Ala Ala
  275      280      285
Tyr Asn Gly Lys Val Asp Leu Ile Ala Thr Gly Phe Gly Glu Ala Pro
  290      295      300
Thr Ala Val Asn Gln Ala Ile Asn Tyr Ile Tyr Pro Asp Arg Asp Asn
  305      310      315      320
Arg Val Val His Ser Thr Ser Leu Ile Asp
      325      330

```

<210> 270

<211> 325

<212> PRT

<213> Sulfolobus solfataricus

<400> 270

```

Met Pro Leu Lys Thr Tyr Asp Thr Ile Ile Val Gly Ala Gly Ile Ala
  1      5      10      15
Gly Leu Ser Ala Ala Leu Tyr Ser Ser Arg Gln Lys Leu Ser Thr Leu
      20      25      30
Val Leu Ser Lys Asp Leu Gly Gly Gln Leu Thr Leu Thr Asp Leu Ile
  35      40      45
Glu Asn Tyr Pro Gly Ile Glu Ser Thr Gly Gly Leu Thr Leu Ala Gln
  50      55      60
Lys Ile Glu Lys Gln Ala Lys Lys Phe Gly Ala Glu Phe Ile Tyr Gly
  65      70      75      80
Glu Glu Val Lys Glu Ile Ala Gln Glu Ser Asp Leu Phe Ile Ile Lys
      85      90      95
Gly Ile Lys Gly Glu Tyr Ala Gly Arg Ala Leu Ile Leu Ala Phe Gly
  100      105      110
Lys Thr Pro Arg Glu Ile Asn Val Pro Gly Glu Gln Glu Phe Lys Gly
  115      120      125
Lys Gly Val Ser Tyr Cys Ala Ile Cys Asp Ala Ala Phe Phe Lys Gly
  130      135      140
Lys Pro Ala Ala Val Ile Gly Glu Gly Glu Pro Gly Ile Glu Ala Ile
  145      150      155      160
Glu Leu Leu Ser Asn Tyr Ala Asn Pro Ala Tyr Tyr Ile Thr Ser Ser

```

Ser	Tyr	Leu	Ala	165	Gly	Glu	Glu	Glu	Ile	170	Val	Lys	Asn	Val	Val	175	Asn	Lys
			180						185							190		
Pro	Thr	Val	Lys	Ile	Leu	Thr	Ser	Ser	Arg	Val	Leu	Glu	Ile	Arg	Gly			
		195					200								205			
Asn	Ser	Lys	Val	Glu	Glu	Leu	Val	Ile	Lys	Arg	Gly	Asp	Glu	Ile	Leu			
	210					215						220						
Gln	Leu	Lys	Val	Asp	Gly	Val	Ile	Ile	Glu	Met	Gly	Tyr	Thr	Leu	Lys			
225					230					235					240			
Thr	Glu	Phe	Leu	Lys	Gly	Phe	Val	Glu	Leu	Asn	Glu	Lys	Gly	Glu	Ile			
				245					250						255			
Ile	Val	Asp	Glu	Leu	Gly	Arg	Thr	Ser	Arg	Glu	Gly	Val	Phe	Ala	Ala			
		260						265						270				
Gly	Asp	Val	Thr	Gln	Thr	Pro	Tyr	Lys	Gln	Ala	Val	Val	Ala	Ala	Ala			
	275						280						285					
Glu	Gly	Val	Lys	Ala	Ala	Leu	Ser	Ala	Tyr	Asn	Tyr	Ile	Arg	Ser	Lys			
	290					295					300							
Arg	Gly	Leu	Pro	Pro	Val	Thr	Val	Asp	Trp	Lys	Ala	Glu	Lys	Lys	Lys			
305					310					315					320			
Val	Ser	Phe	Arg	Leu														
				325														

<210> 271

<211> 323

<212> PRT

<213> Sulfolobus solfataricus

<400> 271

Met	Ser	Leu	Leu	Pro	Arg	Thr	Thr	Ser	Val	Lys	Pro	Gly	Glu	Lys	Phe			
1				5				10						15				
Asp	Val	Ile	Ile	Val	Gly	Leu	Gly	Pro	Ala	Ala	Tyr	Gly	Ala	Ala	Leu			
		20						25					30					
Tyr	Ser	Ala	Arg	Tyr	Met	Leu	Lys	Thr	Leu	Val	Ile	Gly	Glu	Thr	Pro			
	35					40						45						
Gly	Gly	Gln	Leu	Thr	Glu	Ala	Gly	Ile	Val	Asp	Asp	Tyr	Leu	Gly	Leu			
	50					55				60								
Ile	Glu	Ile	Gln	Ala	Ser	Asp	Met	Ile	Lys	Val	Phe	Asn	Lys	His	Ile			
65					70					75					80			
Glu	Lys	Tyr	Glu	Val	Pro	Val	Leu	Leu	Asp	Ile	Val	Glu	Lys	Ile	Glu			
				85					90					95				
Asn	Arg	Gly	Asp	Glu	Phe	Val	Val	Lys	Thr	Lys	Arg	Lys	Gly	Glu	Phe			
		100						105					110					
Lys	Ala	Asp	Ser	Val	Ile	Leu	Gly	Ile	Gly	Val	Lys	Arg	Arg	Lys	Leu			
	115						120					125						
Gly	Val	Pro	Gly	Glu	Gln	Glu	Phe	Ala	Gly	Arg	Gly	Ile	Ser	Tyr	Cys			
	130					135					140							
Ser	Val	Cys	Asp	Ala	Pro	Leu	Phe	Lys	Asn	Arg	Val	Val	Ala	Val	Ile			
145					150					155					160			
Gly	Gly	Gly	Asp	Ser	Ala	Leu	Glu	Gly	Ala	Glu	Ile	Leu	Ser	Ser	Tyr			
			165						170						175			
Ser	Thr	Lys	Val	Tyr	Leu	Ile	His	Arg	Arg	Asp	Thr	Phe	Lys	Ala	Gln			
		180						185					190					
Pro	Ile	Tyr	Val	Glu	Thr	Val	Lys	Lys	Lys	Pro	Asn	Val	Glu	Phe	Val			
	195						200						205					
Leu	Asn	Ser	Val	Val	Lys	Glu	Ile	Lys	Gly	Asp	Lys	Val	Val	Lys	Gln			
	210					215					220							
Val	Val	Val	Glu	Asn	Leu	Lys	Thr	Gly	Glu	Ile	Lys	Glu	Leu	Asn	Val			
225					230					235					240			
Asn	Gly	Val	Phe	Ile	Glu	Ile	Gly	Phe	Asp	Pro	Pro	Thr	Asp	Phe	Ala			
			245						250					255				
Lys	Ser	Asn	Gly	Ile	Glu	Thr	Asp	Thr	Asn	Gly	Tyr	Ile	Lys	Val	Asp			
		260						265					270					
Glu	Trp	Met	Arg	Thr	Ser	Val	Pro	Gly	Val	Phe	Ala	Ala	Gly	Asp	Cys			
	275						280					285						
Thr	Ser	Ala	Trp	Leu	Gly	Phe	Arg	Gln	Val	Ile	Thr	Ala	Val	Ala	Gln			
	290					295					300							

Gly Ala Val Ala Ala Thr Ser Ala Tyr Arg Tyr Val Thr Glu Lys Lys
 305 310 315 320
 Gly Lys Lys

<210> 272
 <211> 332
 <212> PRT
 <213> *Sulfolobus solfataricus*

<400> 272
 Met Asp Glu Tyr Asp Ile Val Val Ile Gly Gly Gly Pro Val Gly Leu
 1 5 10 15
 Phe Gly Thr Phe Tyr Ala Gly Leu Arg Asp Met Lys Thr Leu Leu Ile
 20 25 30
 Asp Ala Gln Asp Glu Leu Gly Gly Gln Leu Val Ser Leu Tyr Pro Glu
 35 40 45
 Lys Ile Val Tyr Asp Val Gly Gly Leu Ala Gly Ile Gln Ala Tyr Glu
 50 55 60
 Leu Ala Gln Arg Leu Ile Glu Gln Ala Lys Met Phe Gly Pro Asp Ile
 65 70 75 80
 Lys Val Asn Glu Leu Ala Asp Met Ile Glu Lys Thr Asn Asp Asn Met
 85 90 95
 Trp Ile Val Lys Thr Asp Lys Ala Thr Tyr Lys Thr Lys Thr Ile Phe
 100 105 110
 Ile Ala Ala Gly Ile Gly Lys Ile Val Pro Ser Arg Leu Gly Ala Lys
 115 120 125
 Gly Glu Ile Glu Tyr Glu Asn Arg Gly Val Tyr Tyr Thr Val Arg Arg
 130 135 140
 Lys Lys Asp Phe Glu Gly Lys Arg Val Leu Ile Val Gly Gly Gly Asp
 145 150 155 160
 Ser Ala Val Asp Trp Ala Leu Thr Leu Ala Pro Val Ala Lys Ser Val
 165 170 175
 Thr Leu Ile His Arg Arg Asp Gln Phe Arg Ala His Glu Arg Ser Val
 180 185 190
 Lys Glu Leu Phe Arg Val Ala Asn Val Tyr Val Trp His Glu Leu Lys
 195 200 205
 Glu Val Lys Gly Asp Gly Asn Lys Val Thr Gln Ala Ile Ile Phe Asp
 210 215 220
 Asn Arg Thr Lys Glu Glu Lys Val Leu Asp Val Asp Ser Val Ile Ile
 225 230 235 240
 Ser Ile Gly Tyr Lys Gly Asp Leu Gly Asn Ile Pro Lys Trp Gly Val
 245 250 255
 Thr Met Lys Gly Arg Asp Ile Val Val Asn Gly Arg Met Glu Thr Asn
 260 265 270
 Leu Pro Gly Val Tyr Ala Gly Gly Asp Ile Val Gln Met Glu Gly Ser
 275 280 285
 Pro Lys Leu Ala Leu Ile Ala Val Gly Phe Ala His Ala Ala Ile Ala
 290 295 300
 Ile Ser Val Ala Lys Lys Tyr Val Glu Pro Asn Ala Ser Leu Phe Ala
 305 310 315 320
 Gly His Ser Ser Glu Met Asp Lys Phe Lys Pro Lys
 325 330

<210> 273
 <211> 324
 <212> PRT
 <213> *Rhizobium loti*

<400> 273
 Met Thr Thr Lys His Ala Pro Val Leu Ile Ile Gly Ser Gly Pro Ala
 1 5 10 15
 Gly Tyr Thr Ala Ala Val Tyr Ala Ala Arg Ala Met Leu Lys Pro Met
 20 25 30
 Leu Val Ala Gly Leu Gln Gln Gly Gly Gln Leu Met Ile Thr Thr Asp

Val	Glu	Asn	Tyr	Pro	Gly	Phe	Ala	Asp	Pro	Ile	Gln	Gly	Pro	Trp	Leu
50	55	60	65	70	75	80	85	90	95	100	105	110	115	120	125
Met	Glu	Gln	Met	Met	Lys	Gln	Ala	Glu	His	Val	Gly	Thr	Asp	Ile	Ile
Asn	Asp	Ile	Ile	Thr	Glu	Val	Asp	Leu	Asn	Val	Arg	Pro	Phe	Arg	Ala
Lys	Gly	Asp	Ser	Gly	Thr	Thr	Tyr	Thr	Ala	Asp	Ala	Leu	Ile	Ile	Ala
Thr	Gly	Ala	Gln	Ala	Lys	Trp	Leu	Gly	Ile	Pro	Thr	Glu	Gln	Asp	Phe
Met	Gly	Phe	Gly	Val	Ser	Ala	Cys	Ala	Thr	Cys	Asp	Gly	Phe	Phe	Tyr
Arg	Gly	Lys	Asp	Val	Ala	Val	Val	Gly	Gly	Gly	Asn	Ser	Ala	Val	Glu
Glu	Ala	Leu	Tyr	Leu	Ser	Asn	Leu	Ala	Lys	Ser	Val	Thr	Val	Ile	His
Arg	Arg	Ser	Asp	Phe	Arg	Ala	Glu	Arg	Ile	Leu	Arg	Glu	Arg	Leu	Leu
Gln	Lys	Asp	Asn	Val	Arg	Val	Ile	Trp	Asp	Thr	Val	Val	Asp	Glu	Ile
Thr	Gly	Arg	Pro	Gly	Lys	Ala	Pro	Leu	Pro	Pro	Ser	Val	Glu	Gly	Leu
Lys	Leu	Lys	His	Ala	Val	Thr	Gly	Ala	Glu	Thr	His	Leu	Lys	Val	Asp
Gly	Val	Phe	Val	Ala	Ile	Gly	His	Ala	Pro	Ala	Val	Glu	Leu	Phe	Val
Gly	Lys	Leu	Lys	Gln	Lys	Pro	Asn	Gly	Tyr	Leu	Trp	Thr	Ala	Pro	Asn
Ser	Thr	Arg	Thr	Asp	Val	Pro	Gly	Val	Phe	Ala	Ala	Gly	Asp	Val	Thr
Asp	Asp	Val	Tyr	Arg	Gln	Ala	Val	Thr	Ala	Ala	Gly	Leu	Gly	Cys	Met
Ala	Ala	Leu	Glu	Ala	Glu	Lys	Tyr	Leu	Ala	Gly	Ile	Glu	Val	His	Arg
Glu	Ala	Ala	Glu												

<210> 274

<211> 343

<212> PRT

<213> Rhizobium loti

<400> 274

Met	Thr	Gly	Ile	Ile	Ser	Thr	Asp	Val	Leu	Ile	Val	Gly	Ala	Gly	Pro
1	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75
Val	Gly	Leu	Phe	Ala	Val	Phe	Glu	Leu	Gly	Leu	Phe	Asp	Met	Lys	Cys
His	Leu	Ile	Asp	Ile	Leu	Asp	Lys	Pro	Gly	Gly	Gln	Cys	Ala	Glu	Leu
Tyr	Pro	Glu	Lys	Pro	Ile	Tyr	Asp	Ile	Pro	Gly	Trp	Pro	Ser	Ile	Ser
Ala	Gln	Gly	Leu	Val	Asp	Lys	Leu	Leu	Glu	Gln	Ile	His	Pro	Phe	Lys
Pro	Asp	Phe	Thr	Tyr	Asn	Arg	Met	Val	Ser	Ser	Leu	Glu	Lys	Leu	Glu
Asp	Gly	Ser	Phe	Arg	Val	Thr	Thr	Asp	Glu	Asn	Glu	Val	Phe	Glu	Ala
Lys	Val	Val	Val	Ile	Ala	Ala	Gly	Gly	Gly	Ser	Phe	Gln	Pro	Lys	Arg
Pro	Pro	Ile	Pro	Gly	Ile	Glu	Pro	Tyr	Glu	Gly	Lys	Ser	Val	Phe	Tyr
Ser	Val	Arg	Arg	Met	Glu	Asp	Phe	Arg	Gly	His	Asp	Leu	Val	Ile	Val
Gly	Gly	Gly	Asp	Ser	Ala	Leu	Asp	Trp	Thr	Leu	Asn	Leu	Gln	Pro	Val

Ala Lys Ser Val Thr Leu Val His Arg Arg Pro Glu Phe Arg Ala Ala
 180 185 190
 Pro Asp Ser Val Asn Lys Met Tyr Ala Met Gln Glu Met Lys Gln Leu
 195 200 205
 Glu Phe Arg Val Gly Gln Val Thr Gly Leu Thr Gly Ala Asp Gly Gln
 210 215 220
 Leu Ser Ser Ala Thr Ile Lys Gly Gly Pro Asp Gly Asp Ile Glu Val
 225 230 235 240
 Pro Cys Thr Arg Met Leu Pro Phe Phe Gly Leu Thr Met Lys Leu Gly
 245 250 255
 Pro Ile Ala Glu Trp Gly Leu Asn Leu His Glu Asn Leu Ile Pro Val
 260 265 270
 Asp Thr Glu Lys Phe Gln Thr Ser Val Pro Gly Ile Phe Ala Val Gly
 275 280 285
 Asp Ile Asn Ser Tyr Pro Gly Lys Leu Lys Leu Ile Leu Ser Gly Phe
 290 295 300
 His Glu Val Ala Leu Met Ala Gln Ala Ala Lys Arg Ile Val Ser Pro
 305 310 315 320
 Gly Glu Arg Ile Val Phe Gln Tyr Thr Thr Ser Ser Thr Ser Leu Gln
 325 330 335
 Lys Lys Leu Gly Val Val Gly
 340

<210> 275
 <211> 15
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<220>
 <221> VARIANT
 <222> 9, 11
 <223> Xaa = Any Amino Acid

<400> 275
 Val His Asn Ile Val Thr Ile Ile Xaa Ser Xaa Pro Ala Ala His
 1 5 10 15

<210> 276
 <211> 104
 <212> PRT
 <213> *Staphylococcus aureus*

<400> 276
 Met Ala Ile Val Lys Val Thr Asp Ala Asp Phe Asp Ser Lys Val Glu
 1 5 10 15
 Ser Gly Val Gln Leu Val Asp Phe Trp Ala Thr Trp Cys Gly Pro Cys
 20 25 30
 Lys Met Ile Ala Pro Val Leu Glu Glu Leu Ala Ala Asp Tyr Glu Gly
 35 40 45
 Lys Ala Asp Ile Leu Lys Leu Asp Val Asp Glu Asn Pro Ser Thr Ala
 50 55 60
 Ala Lys Tyr Glu Val Met Ser Ile Pro Thr Leu Ile Val Phe Lys Asp
 65 70 75 80
 Gly Gln Pro Val Asp Lys Val Val Gly Phe Gln Pro Lys Glu Asn Leu
 85 90 95
 Ala Glu Val Leu Asp Lys His Leu
 100

<210> 277
 <211> 92
 <212> PRT
 <213> *Staphylococcus xylosus*

<400> 277

Met Ala Glu Gln Val Asp Phe Asp Ile Ala Ile Ile Gly Ala Gly Pro
 1 5 10 15
 Ala Gly Met Thr Ala Ala Val Tyr Ala Ser Arg Ala Asn Leu Ser Thr
 20 25 30
 Val Met Ile Glu Arg Gly Met Pro Gly Gly Gln Met Ala Asn Thr Glu
 35 40 45
 Glu Val Glu Asn Phe Pro Gly Phe Glu Met Val Thr Gly Pro Asp Leu
 50 55 60
 Ser Thr Lys Met Phe Glu His Ala Lys Lys Phe Gly Ala Lys Tyr Gln
 65 70 75 80
 Tyr Gly Asp Ile Lys Ser Ile Glu Asp Lys Gly Ser
 85 90

<210> 278

<211> 319

<212> PRT

<213> Thermoplasma acidophilum

<400> 278

Met Glu Phe Asn Leu His Ala Val Ser Ser Glu Glu Lys Glu Arg Asp
 1 5 10 15
 Phe Asp Val Val Ile Val Gly Ala Gly Ala Ala Gly Phe Ser Ala Ala
 20 25 30
 Val Tyr Ala Ala Arg Ser Gly Phe Ser Val Ala Ile Leu Asp Lys Ala
 35 40 45
 Val Ala Gly Gly Leu Thr Ala Glu Ala Pro Leu Val Glu Asn Tyr Leu
 50 55 60
 Gly Phe Lys Ser Ile Val Gly Ser Glu Leu Ala Lys Leu Phe Ala Asp
 65 70 75 80
 His Ala Ala Asn Tyr Ala Lys Ile Arg Glu Gly Val Glu Val Arg Ser
 85 90 95
 Ile Lys Lys Thr Gln Gly Gly Phe Asp Ile Glu Thr Asn Asp Asp Thr
 100 105 110
 Tyr His Ala Lys Tyr Val Ile Ile Thr Thr Gly Thr Thr His Lys His
 115 120 125
 Leu Gly Val Lys Gly Glu Ser Glu Tyr Phe Gly Lys Gly Thr Ser Tyr
 130 135 140
 Cys Ser Thr Cys Asp Gly Tyr Leu Phe Lys Gly Lys Arg Val Val Thr
 145 150 155 160
 Ile Gly Gly Gly Asn Ser Gly Ala Ile Ala Ala Ile Ser Met Ser Glu
 165 170 175
 Tyr Val Lys Asn Val Thr Ile Ile Glu Tyr Met Pro Lys Tyr Met Cys
 180 185 190
 Glu Asn Ala Tyr Val Gln Glu Ile Lys Lys Arg Asn Ile Pro Tyr Ile
 195 200 205
 Met Asn Ala Gln Val Thr Glu Ile Val Gly Asp Gly Lys Lys Val Thr
 210 215 220
 Gly Val Lys Tyr Lys Asp Arg Thr Thr Gly Glu Glu Lys Leu Ile Glu
 225 230 235 240
 Thr Asp Gly Val Phe Ile Tyr Val Gly Leu Ile Pro Gln Thr Ser Phe
 245 250 255
 Leu Lys Asp Ser Gly Val Lys Leu Asp Glu Arg Gly Tyr Ile Val Val
 260 265 270
 Asp Ser Arg Gln Arg Thr Ser Val Pro Gly Val Tyr Ala Ala Gly Asp
 275 280 285
 Val Thr Ser Gly Asn Phe Ala Gln Ile Ala Ser Ala Val Gly Asp Gly
 290 295 300
 Cys Lys Ala Ala Leu Ser Leu Tyr Ser Asp Ser Ile Ser Lys Lys
 305 310 315

<210> 279

<211> 317

<212> PRT

<213> Thermotoga maritima

<400> 279
 Met Val Phe Phe Asp Thr Gly Ser Leu Lys Lys Lys Glu Ile Lys Asp
 1 5 10 15
 Lys Tyr Asp Ile Val Val Val Gly Gly Gly Pro Ala Gly Leu Thr Ser
 20 25 30
 Ala Ile Tyr Ala Arg Arg Ala Gly Leu Ser Val Leu Val Val Glu Lys
 35 40 45
 Ala Ile Glu Gly Gly Tyr Val Asn Leu Thr His Leu Val Glu Asn Tyr
 50 55 60
 Pro Gly Phe Pro Ala Ile Ser Gly Glu Glu Leu Ala Ser Lys Phe Lys
 65 70 75 80
 Glu His Ala Glu Lys Phe Gly Ala Asp Ile Tyr Asn Ala Glu Val Val
 85 90 95
 Lys Leu Glu Val Gln Gly Asp Lys Lys Val Val Glu Leu Asp Asp Gly
 100 105 110
 Lys Arg Ile Glu Ala Pro Val Val Ile Val Ala Thr Gly Ala Asn Pro
 115 120 125
 Lys Lys Leu Asn Val Pro Gly Glu Lys Glu Phe Phe Gly Lys Gly Val
 130 135 140
 Ser Tyr Cys Ala Thr Cys Asp Gly Tyr Leu Phe Ala Gly Lys Asp Val
 145 150 155 160
 Ile Val Val Gly Gly Gly Asp Ser Ala Cys Asp Glu Ser Ile Phe Leu
 165 170 175
 Ser Asn Ile Val Asn Lys Ile Thr Met Ile Gln Leu Leu Glu Thr Leu
 180 185 190
 Thr Ala Ala Lys Val Leu Gln Glu Arg Val Leu Asn Asn Pro Lys Ile
 195 200 205
 Glu Val Ile Tyr Asn Ser Thr Val Arg Glu Ile Arg Gly Lys Asp Lys
 210 215 220
 Val Glu Glu Val Val Ile Glu Asn Val Lys Thr Gly Glu Thr Lys Val
 225 230 235 240
 Leu Lys Ala Asp Gly Val Phe Ile Phe Ile Gly Leu Asp Pro Asn Ser
 245 250 255
 Lys Leu Leu Glu Gly Leu Val Glu Leu Asp Pro Tyr Gly Tyr Val Ile
 260 265 270
 Thr Asp Glu Asn Met Glu Thr Ser Val Lys Gly Ile Tyr Ala Val Gly
 275 280 285
 Asp Val Arg Lys Lys Asn Leu Arg Gln Ile Val Thr Ala Val Ala Asp
 290 295 300
 Gly Ala Ile Ala Val Glu His Ala Ala Lys His Tyr Phe
 305 310 315

<210> 280

<211> 326

<212> PRT

<213> Thermoplasma volcanium

<400> 280

Met Asn Leu Tyr Arg Gly Met Glu Phe Asn Leu Arg Ser Val Ser Thr
 1 5 10 15
 Glu Ala Lys Glu Arg Asp Phe Asp Val Ile Ile Ile Gly Ala Gly Ala
 20 25 30
 Ala Gly Phe Ser Ala Ala Val Tyr Ala Ser Arg Ser Gly Leu Ser Ala
 35 40 45
 Val Ile Leu Asp Lys Asn Val Ala Gly Gly Leu Thr Ala Glu Ala Pro
 50 55 60
 Leu Val Glu Asn Tyr Leu Gly Phe Lys Ser Ile Val Gly Ser Asp Leu
 65 70 75 80
 Ala Lys Asn Phe Ala Glu His Ala Ser Glu Tyr Ala Ser Ile Arg Glu
 85 90 95
 Gly Val Glu Val Lys Ser Val Lys Lys Gly Asp Gly Gly Phe Ile Val
 100 105 110
 Asp Thr Ser Asp Gly Glu Tyr His Ser Lys Tyr Ile Ile Ile Thr Thr
 115 120 125
 Gly Thr Thr His Lys His Leu Gly Val Lys Gly Glu Ala Glu Tyr Phe
 130 135 140

Gly Lys Gly Val Ser Tyr Cys Ser Thr Cys Asp Gly Tyr Leu Phe Lys
 145 150 155 160
 Asn Lys Asn Val Val Thr Ile Gly Gly Gly Asn Ser Gly Ala Ile Ala
 165 170 175
 Ala Ile Ser Met Ser Glu Tyr Val Lys Asn Ala Thr Ile Val Glu Tyr
 180 185 190
 Met Pro Arg Tyr Met Cys Glu Asn Ala Tyr Ile Glu Glu Ile Lys Lys
 195 200 205
 Arg Lys Ile Pro Tyr Ile Met Asn Ala Gln Val Thr Glu Ile Val Gly
 210 215 220
 Asp Gly Lys Lys Val Thr Gly Val Lys Tyr Lys Asp Arg Ser Ser Gly
 225 230 235 240
 Glu Glu Lys Thr Leu Pro Ala Asp Gly Val Phe Val Tyr Val Gly Leu
 245 250 255
 Ile Pro Gln Thr Ser Phe Leu Lys Asp Ser Gly Val Lys Leu Asp Glu
 260 265 270
 Arg Gly Tyr Ile Ile Val Asp Gly Arg Gln Arg Thr Asn Val Pro Gly
 275 280 285
 Ile Tyr Ala Ala Gly Asp Val Thr Ser Gly Ser Phe Ala Gln Ile Ala
 290 295 300
 Ser Ala Val Gly Asp Gly Cys Lys Ala Ala Leu Ser Leu Tyr Ser Asp
 305 310 315 320
 Thr Ile Ser Ser Lys Lys
 325

<210> 281

<211> 309

<212> PRT

<213> Ureaplasma parvum

<400> 281

Met Asn Gln Glu Val Tyr Asp Leu Val Ile Ile Gly Ala Gly Pro Ala
 1 5 10 15
 Gly Leu Ala Ala Ala Val Tyr Ala Lys Arg Ser Gly Leu Asn Val Ile
 20 25 30
 Ile Val Glu Lys Gln Phe Pro Gly Gly Lys Ile Ala Leu Thr Ser Asn
 35 40 45
 Val Glu Asn Tyr Leu Gly Ile Asn Ser Ile Pro Gly Pro Glu Leu Ala
 50 55 60
 Tyr Lys Met Tyr Glu Gln Val Leu Asn Leu Asn Val Ser Ile Ile Tyr
 65 70 75 80
 Glu Ala Ala Asp Glu Ile Ser Leu Lys Glu Lys Tyr Lys Lys Ile Lys
 85 90 95
 Leu Thr Thr Gln Thr Leu Ile Thr Lys Thr Val Ile Ile Ala Thr Gly
 100 105 110
 Thr Glu Asn Arg Arg Leu Asn Ile Leu Gly Glu Leu Glu Phe Glu Asn
 115 120 125
 Lys Gly Ile Ser Tyr Cys Ala Ile Cys Asp Gly Pro Leu Tyr Lys Asn
 130 135 140
 Lys Ala Val Ser Val Ile Gly Ser Gly Asn Ser Ala Val Glu Glu Ala
 145 150 155 160
 Ile Tyr Leu Ala Thr Ile Ala Lys Glu Val His Leu Ile Ala Asn Lys
 165 170 175
 Pro Gln Phe Lys Ala Glu Gln Gln Leu Val Gln Ile Ala Asn Asn Thr
 180 185 190
 Pro Asn Ile Lys Ile Tyr Tyr Asn Lys Gln Thr Phe Glu Phe Phe Gly
 195 200 205
 His Gln Phe Leu Glu Gly Leu Lys Phe Arg Asp Leu Ile Thr Asn Glu
 210 215 220
 Val Thr Thr Leu Asn Ile Glu Ala Asn Phe Thr Phe Ile Gly Leu Leu
 225 230 235 240
 Pro Ser Arg Ile Asn Thr Asn Asn Leu Cys Ile Phe Asn Glu Val Asn
 245 250 255
 Gly Phe Ile Thr Thr Asp Lys Asn Met Gln Thr Ser Val Cys Gly Ile
 260 265 270
 Phe Ala Ala Gly Asp Ile Val Asp Lys Asn Val Arg Gln Ile Ala Thr

275 280 285
 Ala Thr Asn Asp Gly Val Ile Ala Ala Leu Tyr Ala Lys Glu Tyr Ile
 290 295 300
 Thr Arg Asn Asn Trp
 305

<210> 282
 <211> 318
 <212> PRT
 <213> *Vibrio cholerae*

<400> 282
 Met Ser Asn Val Lys His Ser Lys Leu Leu Ile Leu Gly Ser Gly Pro
 1 5 10 15
 Ala Gly Tyr Thr Ala Ala Val Tyr Ala Ala Arg Ala Asn Leu Lys Pro
 20 25 30
 Val Leu Val Thr Gly Met Gln Gln Gly Gly Gln Leu Thr Thr Thr Thr
 35 40 45
 Glu Val Glu Asn Trp Pro Gly Asp Ala Glu Gly Leu Thr Gly Pro Ala
 50 55 60
 Leu Met Glu Arg Met Lys Glu His Ala Glu Arg Phe Asp Thr Glu Ile
 65 70 75 80
 Val Phe Asp His Ile Asn Ser Val Asp Leu Ser Ser Arg Pro Phe Arg
 85 90 95
 Leu Thr Gly Asp Ser Gln Glu Tyr Thr Cys Asp Ala Leu Ile Ile Ser
 100 105 110
 Thr Gly Ala Ser Ala Lys Tyr Leu Gly Leu Glu Ser Glu Glu Ala Phe
 115 120 125
 Lys Gly Arg Gly Val Ser Ala Cys Ala Thr Cys Asp Gly Phe Phe Tyr
 130 135 140
 Arg Asn Gln Lys Val Ala Val Val Gly Gly Gly Asn Thr Ala Val Glu
 145 150 155 160
 Glu Ala Leu Tyr Leu Ser Asn Ile Ala Ser Glu Val His Leu Val His
 165 170 175
 Arg Arg Asp Ser Phe Arg Ser Glu Lys Ile Leu Ile Asp Arg Leu Met
 180 185 190
 Asp Lys Val Ala Asn Gly Asn Ile Val Leu His Thr His Arg Thr Leu
 195 200 205
 Asp Glu Val Leu Gly Asp Glu Met Gly Val Thr Gly Val Arg Leu Lys
 210 215 220
 Asp Thr Gln Ser Asp Met Thr Glu Asn Leu Asp Val Met Gly Val Phe
 225 230 235 240
 Ile Ala Ile Gly His Gln Pro Asn Ser Gln Ile Phe Glu Gly Gln Leu
 245 250 255
 Glu Met Lys Asn Gly Tyr Ile Val Val Lys Ser Gly Leu Glu Gly Asn
 260 265 270
 Ala Thr Gln Thr Ser Ile Glu Gly Val Phe Ala Ala Gly Asp Val Met
 275 280 285
 Asp His Asn Tyr Arg Gln Ala Ile Thr Ser Ala Gly Thr Gly Cys Met
 290 295 300
 Ala Ala Leu Asp Ala Glu Arg Tyr Leu Asp Ser Gln Gly Lys
 305 310 315

<210> 283
 <211> 321
 <212> PRT
 <213> *Xylella fastidiosa*

<400> 283
 Met Ser Asp Tyr Pro Ala Ser Ala Lys His Ser Arg Leu Leu Ile Leu
 1 5 10 15
 Gly Ser Gly Pro Ala Gly Trp Thr Ala Ala Val Tyr Ala Ala Arg Ala
 20 25 30
 Asn Leu Gln Pro Val Leu Ile Thr Gly Leu Gln Gln Gly Gly Gln Leu
 35 40 45

```

Met Thr Thr Thr Glu Val Asp Asn Trp Pro Gly Asp Ala His Gly Leu
 50      55      60
Met Gly Pro Asp Leu Met Glu Arg Met Gln Ala His Ala Glu Arg Phe
65      70      75      80
Asp Thr Lys Val Ile Phe Asp Gln Ile Tyr Lys Ala Asp Leu Ser Thr
 85      90      95
Arg Pro Phe Thr Leu Phe Gly Asp Ser Gly Leu Tyr Thr Cys Asp Gly
100      105      110
Leu Ile Ile Ala Thr Gly Ala Asn Ala Lys Tyr Leu Gly Ile Pro Ser
115      120      125
Glu Glu Ala Phe Lys Gly Arg Gly Val Ser Ala Cys Ala Thr Cys Asp
130      135      140
Gly Phe Phe Tyr Arg Asp Gln Asp Val Ala Val Ile Gly Gly Gly Asn
145      150      155      160
Thr Ala Val Glu Glu Ala Leu Tyr Leu Ser Asn Ile Ala Arg Lys Val
165      170      175
Tyr Leu Ile His Arg Arg Asp Lys Leu Arg Ala Glu Lys Ile Met Gln
180      185      190
Asn Lys Leu Phe Ser Lys Ala Ala Thr Gly Lys Ile Glu Leu Ile Trp
195      200      205
Asn Asn Ala Val Glu Glu Val Leu Gly Asn Asp Ala Ser Val Thr Gly
210      215      220
Val Arg Ile Arg Ser Thr Gln Asp Ser Ser Thr Arg Asp Ile Asp Val
225      230      235      240
Gln Gly Leu Phe Val Ala Ile Gly His His Pro Asn Thr Asp Leu Phe
245      250      255
Ala Gly Gln Leu Ala Met Asn Asn Gly Tyr Leu Gln Ile His Ser Gly
260      265      270
Thr Ala Gly Asn Val Thr Gln Thr Ser Val Glu Gly Val Phe Ala Ala
275      280      285
Gly Asp Val Ala Asp Gln His Tyr Arg Gln Ala Ile Thr Ser Ala Gly
290      295      300
Phe Gly Cys Met Ala Ala Leu Asp Ala Glu Arg Phe Leu Asp Lys Gly
305      310      315      320
Asn

```

<210> 284

<211> 318

<212> PRT

<213> *Zymomonas mobilis*

<400> 284

```

Met Ser Ala Asp Pro Ile Ser Thr Arg Val Phe Ile Leu Gly Ser Gly
 1      5      10      15
Pro Ala Gly Leu Thr Ala Ala Ile Tyr Ala Ala Arg Ala Gly Leu Asn
20      25      30
Pro Ile Val Ala Gln Gly Leu Gln Pro Gly Gly Gln Leu Thr Ile Thr
35      40      45
Thr Glu Val Glu Asn Phe Pro Gly Phe Arg Glu Pro Ile Gln Gly Pro
50      55      60
Trp Leu Met Glu Glu Met Gln Ala Gln Ala Glu Asn Val Gly Ala Lys
65      70      75      80
Leu Val Trp Asp Ile Ile Thr Ser Val Asp Phe Ser Gln Arg Pro Tyr
85      90      95
Arg Leu Met Gly Asp Gly Gly Gln Val Tyr Leu Ala Asp Ser Leu Ile
100      105      110
Ile Ser Thr Gly Ala Gln Ala Arg Trp Leu Gly Leu Glu Ser Glu Thr
115      120      125
Ala Leu Arg Gly Lys Gly Ile Ser Ala Cys Ala Thr Cys Asp Gly Phe
130      135      140
Phe Phe Arg Gly Lys Lys Val Val Val Ile Gly Gly Asn Thr Ala
145      150      155      160
Val Glu Glu Ala Leu Tyr Leu Thr Asn His Ser Pro Glu Val Thr Leu
165      170      175
Ile His Arg Arg Asp Ser Leu Arg Ala Glu Lys Ile Met Gln Lys Arg

```

```

      180      185      190
Leu Leu Ala Asn Pro Lys Ile Lys Ile Arg Trp Asn Ser Glu Val Ala
      195      200      205
Glu Phe Ile Ala Gly Glu Asp Ser Ala Leu Ser Ala Val Lys Leu Lys
      210      215      220
Asp Thr Lys Thr Gly Glu Glu Ser Leu Leu Glu Thr Glu Gly Ala Phe
      225      230      235
Ile Ala Ile Gly His Lys Pro Ala Thr Glu Leu Phe Gln Gly His Leu
      245      250      255
Lys Leu Asp Asp Glu Gly Tyr Ile Glu Val Thr Pro Gly Thr Thr Gln
      260      265      270
Thr Ser Ile Lys Gly Ile Phe Ala Cys Gly Asp Val Met Asp Lys His
      275      280      285
Tyr Arg Gln Ala Val Thr Ala Ala Gly Thr Gly Cys Met Ala Ala Leu
      290      295      300
Glu Ala Glu Arg Phe Leu Gly Glu Ile Asp Phe Lys Glu Asp
      305      310      315

```

<210> 285
 <211> 122
 <212> PRT
 <213> Bos taurus

```

<400> 285
Lys Leu Met His Gln Ala Ala Leu Leu Gly Gln Ala Leu Thr Asp Ser
  1      5      10      15
Arg Lys Phe Gly Trp Glu Tyr Ser Gln Val Arg His Ser Trp Ala
      20      25      30
Thr Met Thr Glu Ala Ile Gln Ser His Ile Gly Ser Leu Ser Trp Gly
      35      40      45
His Arg Leu Ala Leu Arg Glu Lys Ala Val Thr Tyr Val Asn Ser Phe
      50      55      60
Gly Glu Phe Val Glu His Lys Val Lys Ala Thr Asn Glu Lys Gly
      65      70      75      80
Gln Glu Val Leu Tyr Thr Ala Ala Lys Phe Val Ile Ala Thr Gly Glu
      85      90      95
Arg Pro Arg Tyr Leu Gly Ile Pro Gly Asp Arg Glu Tyr Cys Ile Thr
      100      105      110
Ser Asp Asp Leu Phe Ser Leu Pro Tyr Cys
      115      120

```

<210> 286
 <211> 511
 <212> PRT
 <213> Bos taurus

```

<400> 286
Met Ala Ala Leu Arg Gly Ala Ala Ala Arg Phe Arg Gly Arg Ala Pro
  1      5      10      15
Gly Gly Ala Arg Gly Ala Ala Gly Arg Gln Cys Tyr Asp Leu Leu Val
      20      25      30
Ile Gly Gly Gly Ser Gly Gly Leu Ala Cys Ala Lys Glu Ala Ala Gln
      35      40      45
Leu Gly Lys Lys Val Ala Val Leu Asp Tyr Val Glu Pro Ser Pro Gln
      50      55      60
Gly Thr Arg Trp Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys Ile
      65      70      75      80
Pro Lys Lys Leu Met His Gln Ala Ala Leu Leu Gly Gly Met Ile Arg
      85      90      95
Asp Ala Pro His Tyr Gly Trp Gly Val Ala Gln Ala Pro His Ser Trp
      100      105      110
Ala Thr Leu Ala Asp Ala Val Gln Asn His Val Lys Ser Leu Asn Trp
      115      120      125
Gly His Arg Ile Gln Leu Gln Asp Arg Lys Val Lys Tyr Phe Asn Val
      130      135      140

```

Lys Ala Ser Phe Val Asp Thr His Thr Val Cys Gly Val Ser Lys Gly
 145 150 155 160
 Gly Glu Glu Thr Leu Leu Ser Ala Glu His Ile Val Ile Ala Thr Gly
 165 170 175
 Gly Arg Pro Arg Tyr Pro Thr His Ile Glu Gly Ala Leu Glu Tyr Gly
 180 185 190
 Ile Thr Ser Asp Asp Leu Phe Trp Leu Lys Glu Ser Pro Gly Lys Thr
 195 200 205
 Leu Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Leu Leu
 210 215 220
 Thr Gly Leu Gly Leu Asp Thr Thr Val Met Ile Arg Ser Val Pro Leu
 225 230 235 240
 Arg Ala Phe Asp Gln Met Ala Ser Leu Val Thr Glu His Met Ala
 245 250 255
 Gly His Gly Thr Arg Ile Leu Arg Gly Cys Ala Pro Glu Lys Val Glu
 260 265 270
 Lys Leu Pro Gly Gln Gln Leu Arg Val Thr Trp Val Asp Leu Thr Ser
 275 280 285
 Asp Arg Lys Asp Ala Gly Thr Phe Asp Thr Val Leu Trp Ala Ile Gly
 290 295 300
 Arg Val Pro Glu Thr Ala Ser Leu Asn Leu Glu Lys Ala Gly Val His
 305 310 315 320
 Thr Asn Pro Val Thr Gly Lys Ile Leu Val Asp Ala Gln Glu Thr Thr
 325 330 335
 Ser Val Pro His Ile Tyr Ala Ile Gly Asp Val Ala Glu Gly Arg Pro
 340 345 350
 Glu Leu Thr Pro Thr Ala Ile Met Ala Gly Arg Leu Leu Ala Gln Arg
 355 360 365
 Leu Ser Gly Arg Thr Ser Asp Leu Met Asp Tyr Ser Ser Val Pro Thr
 370 375 380
 Thr Val Phe Thr Pro Leu Glu Tyr Gly Cys Val Gly Leu Ser Glu Glu
 385 390 395 400
 Ala Ala Val Ala Arg His Gly Glu Glu His Val Glu Val Tyr His Ala
 405 410 415
 Phe Tyr Lys Pro Leu Glu Phe Thr Val Pro Gln Arg Asp Ala Ser Gln
 420 425 430
 Cys Tyr Ile Lys Met Val Cys Leu Arg Glu Pro Pro Gln Leu Val Leu
 435 440 445
 Gly Leu His Phe Leu Gly Pro Asn Ala Gly Glu Val Ile Gln Gly Phe
 450 455 460
 Ala Leu Gly Ile Lys Cys Gly Ala Ser Tyr Gln Gln Leu Met Arg Thr
 465 470 475 480
 Val Gly Ile His Pro Thr Cys Ala Glu Glu Val Ala Lys Leu Arg Ile
 485 490 495
 Ser Lys Arg Ser Gly Leu Asp Pro Thr Val Thr Gly Cys Cys Gly
 500 505 510

<210> 287

<211> 525

<212> PRT

<213> Caenorhabditis elegans

<220>

<221> VARIANT

<222> 524

<223> Xaa = Any Amino Acid

<400> 287

Met Tyr Ile Lys Gly Asn Ala Val Gly Gly Leu Lys Glu Leu Lys Ala
 1 5 10 15
 Leu Lys Gln Asp Tyr Leu Lys Glu Trp Leu Arg Asp His Thr Tyr Asp
 20 25 30
 Leu Ile Val Ile Gly Gly Gly Ser Gly Gly Leu Ala Ala Lys Glu
 35 40 45
 Ala Ser Arg Leu Gly Lys Lys Val Ala Cys Leu Asp Phe Val Lys Pro
 50 55 60

Ser Pro Gln Gly Thr Ser Trp Gly Leu Gly Gly Thr Cys Val Asn Val
 65 70 75 80
 Gly Cys Ile Pro Lys Lys Leu Met His Gln Ala Ser Leu Leu Gly His
 85 90 95
 Ser Ile His Asp Ala Lys Lys Tyr Gly Trp Lys Leu Pro Glu Gly Lys
 100 105 110
 Val Glu His Gln Trp Asn His Leu Arg Asp Ser Val Gln Asp His Ile
 115 120 125
 Ala Ser Leu Asn Trp Gly Tyr Arg Val Gln Leu Arg Glu Lys Thr Val
 130 135 140
 Thr Tyr Ile Asn Ser Tyr Gly Glu Phe Thr Gly Pro Phe Glu Ile Ser
 145 150 155 160
 Ala Thr Asn Lys Lys Lys Lys Val Glu Lys Leu Thr Ala Asp Arg Phe
 165 170 175
 Leu Ile Ser Thr Gly Leu Arg Pro Lys Tyr Pro Glu Ile Pro Gly Val
 180 185 190
 Lys Glu Tyr Thr Ile Thr Ser Asp Asp Leu Phe Gln Leu Pro Tyr Ser
 195 200 205
 Pro Gly Lys Thr Leu Cys Val Gly Ala Ser Tyr Val Ser Leu Glu Cys
 210 215 220
 Ala Gly Phe Leu His Gly Phe Gly Phe Asp Val Thr Val Met Val Arg
 225 230 235 240
 Ser Ile Leu Leu Arg Gly Phe Asp Gln Asp Met Ala Glu Arg Ile Arg
 245 250 255
 Lys His Met Ile Ala Tyr Gly Met Lys Phe Glu Ala Gly Val Pro Thr
 260 265 270
 Arg Ile Glu Gln Ile Asp Glu Lys Thr Asp Glu Lys Ala Gly Lys Tyr
 275 280 285
 Arg Val Phe Trp Pro Lys Lys Asn Glu Glu Thr Gly Glu Met Gln Glu
 290 295 300
 Val Ser Glu Glu Tyr Asn Thr Ile Leu Met Ala Ile Gly Arg Glu Ala
 305 310 315 320
 Val Thr Asp Asp Val Gly Leu Thr Thr Ile Gly Val Glu Arg Ala Lys
 325 330 335
 Ser Lys Lys Val Leu Gly Arg Arg Glu Gln Ser Thr Thr Ile Pro Trp
 340 345 350
 Val Tyr Ala Ile Gly Asp Val Leu Glu Gly Thr Pro Glu Leu Thr Pro
 355 360 365
 Val Ala Ile Gln Ala Gly Arg Val Leu Met Arg Arg Ile Phe Asp Gly
 370 375 380
 Ala Asn Glu Leu Thr Glu Tyr Asp Gln Ile Pro Thr Thr Val Phe Thr
 385 390 395 400
 Pro Leu Glu Tyr Gly Cys Cys Gly Leu Ser Glu Glu Asp Ala Met Met
 405 410 415
 Lys Tyr Gly Lys Asp Asn Ile Ile Ile Tyr His Asn Val Phe Asn Pro
 420 425 430
 Leu Glu Tyr Thr Ile Ser Glu Arg Met Asp Lys Asp His Cys Tyr Leu
 435 440 445
 Lys Met Ile Cys Leu Arg Asn Glu Glu Glu Lys Val Val Gly Phe His
 450 455 460
 Ile Leu Thr Pro Asn Ala Gly Glu Val Thr Gln Gly Phe Gly Ile Ala
 465 470 475 480
 Leu Lys Leu Ala Ala Lys Lys Ala Asp Phe Asp Arg Leu Ile Gly Ile
 485 490 495
 His Pro Thr Val Ala Glu Asn Phe Thr Thr Leu Thr Leu Glu Lys Lys
 500 505 510
 Glu Gly Asp Glu Glu Leu Gln Ala Ser Gly Cys Xaa Gly
 515 520 525

<210> 288

<211> 667

<212> PRT

<213> Caenorhabditis elegans

<220>

<221> VARIANT

<222> 666

<223> Xaa = Any Amino Acid

<400> 288

```

Met Lys Ser Leu Thr Glu Leu Phe Gly Cys Phe Lys Arg Gln Pro Arg
 1      5      10      15
Gln Gln Glu Ala Ser Ser Pro Ala Asn Pro His Val Ser Asp Thr Leu
 20      25      30
Ser Met Gly Val Ala Ala Ser Gly Met Pro Pro Pro Lys Arg Pro Ala
 35      40      45
Pro Ala Glu Ser Pro Thr Leu Pro Gly Glu Thr Leu Val Asp Ala Pro
 50      55      60
Gly Ile Pro Leu Lys Glu Ala Leu Lys Glu Ala Ala Asn Ser Lys Ile
 65      70      75
Val Ile Phe Tyr Asn Ser Ser Asp Glu Glu Lys Gln Leu Val Glu Phe
 85      90      95
Glu Thr Tyr Leu Asn Ser Leu Lys Glu Pro Ala Asp Ala Glu Lys Pro
100      105      110
Leu Glu Ile Pro Glu Ile Lys Lys Leu Gln Val Ser Arg Ala Ser Gln
115      120      125
Lys Val Ile Gln Tyr Leu Thr Leu His Thr Ser Trp Pro Leu Met Tyr
130      135      140
Ile Lys Gly Asn Ala Val Gly Gly Leu Lys Glu Leu Lys Ala Leu Lys
145      150      155
Gln Asp Tyr Leu Lys Glu Trp Leu Arg Asp His Thr Tyr Asp Leu Ile
165      170      175
Val Ile Gly Gly Gly Ser Gly Gly Leu Ala Ala Ala Lys Glu Ala Ser
180      185      190
Arg Leu Gly Lys Lys Val Ala Cys Leu Asp Phe Val Lys Pro Ser Pro
195      200      205
Gln Gly Thr Ser Trp Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys
210      215      220
Ile Pro Lys Lys Leu Met His Gln Ala Ser Leu Leu Gly His Ser Ile
225      230      235
His Asp Ala Lys Lys Tyr Gly Trp Lys Leu Pro Glu Gly Lys Val Glu
245      250      255
His Gln Trp Asn His Leu Arg Asp Ser Val Gln Asp His Ile Ala Ser
260      265      270
Leu Asn Trp Gly Tyr Arg Val Gln Leu Arg Glu Lys Thr Val Thr Tyr
275      280      285
Ile Asn Ser Tyr Gly Glu Phe Thr Gly Pro Phe Glu Ile Ser Ala Thr
290      295      300
Asn Lys Lys Lys Lys Val Glu Lys Leu Thr Ala Asp Arg Phe Leu Ile
305      310      315
Ser Thr Gly Leu Arg Pro Lys Tyr Pro Glu Ile Pro Gly Val Lys Glu
325      330      335
Tyr Thr Ile Thr Ser Asp Asp Leu Phe Gln Leu Pro Tyr Ser Pro Gly
340      345      350
Lys Thr Leu Cys Val Gly Ala Ser Tyr Val Ser Leu Glu Cys Ala Gly
355      360      365
Phe Leu His Gly Phe Gly Phe Asp Val Thr Val Met Val Arg Ser Ile
370      375      380
Leu Leu Arg Gly Phe Asp Gln Asp Met Ala Glu Arg Ile Arg Lys His
385      390      395
Met Ile Ala Tyr Gly Met Lys Phe Glu Ala Gly Val Pro Thr Arg Ile
405      410      415
Glu Gln Ile Asp Glu Lys Thr Asp Glu Lys Ala Gly Lys Tyr Arg Val
420      425      430
Phe Trp Pro Lys Lys Asn Glu Glu Thr Gly Glu Met Gln Glu Val Ser
435      440      445
Glu Glu Tyr Asn Thr Ile Leu Met Ala Ile Gly Arg Glu Ala Val Thr
450      455      460
Asp Asp Val Gly Leu Thr Thr Ile Gly Val Glu Arg Ala Lys Ser Lys
465      470      475
Lys Val Leu Gly Arg Glu Glu Gln Ser Thr Thr Ile Pro Trp Val Tyr
485      490      495
Ala Ile Gly Asp Val Leu Glu Gly Thr Pro Glu Leu Thr Pro Val Ala

```

```

      500      505      510
Ile Gln Ala Gly Arg Val Leu Met Arg Arg Ile Phe Asp Gly Ala Asn
      515      520      525
Glu Leu Thr Glu Tyr Asp Gln Ile Pro Thr Thr Val Phe Thr Pro Leu
      530      535      540
Glu Tyr Gly Cys Cys Gly Leu Ser Glu Glu Asp Ala Met Met Lys Tyr
      545      550      555
Gly Lys Asp Asn Ile Ile Tyr His Asn Val Phe Asn Pro Leu Glu
      565      570      575
Tyr Thr Ile Ser Glu Arg Met Asp Lys Asp His Cys Tyr Leu Lys Met
      580      585      590
Ile Cys Leu Arg Asn Glu Glu Glu Lys Val Val Gly Phe His Ile Leu
      595      600      605
Thr Pro Asn Ala Gly Glu Val Thr Gln Gly Phe Gly Ile Ala Leu Lys
      610      615      620
Leu Ala Ala Lys Lys Ala Asp Phe Asp Arg Leu Ile Gly Ile His Pro
      625      630      635
Thr Val Ala Glu Asn Phe Thr Thr Leu Thr Leu Glu Lys Lys Glu Gly
      645      650      655
Asp Glu Glu Leu Gln Ala Ser Gly Cys Xaa Gly
      660      665

```

<210> 289

<211> 516

<212> PRT

<213> *Drosophila melanogaster*

<400> 289

```

Met Ser Thr Ile Lys Phe Leu Arg Ser Ser Thr His Asn Ala Leu Arg
  1      5      10      15
Ser Ser Leu Gly Trp Cys Arg Leu Ala Ala Ser Arg Pro Arg Tyr Asp
      20      25      30
Tyr Asp Leu Val Val Leu Gly Gly Ser Ala Gly Leu Ala Cys Ala
      35      40      45
Lys Glu Ala Ala Gly Cys Gly Ala Arg Val Leu Cys Phe Asp Tyr Val
      50      55      60
Lys Pro Thr Pro Val Gly Thr Lys Trp Gly Ile Gly Gly Thr Cys Val
      65      70      75      80
Asn Val Gly Cys Ile Pro Lys Lys Leu Met His Gln Ala Ser Leu Leu
      85      90      95
Gly Glu Ala Val His Glu Ala Val Ala Tyr Gly Trp Asn Val Asp Asp
      100      105      110
Thr Asn Ile Arg Pro Asp Trp Arg Lys Leu Val Arg Ser Val Gln Asn
      115      120      125
His Ile Lys Ser Val Asn Trp Val Thr Arg Val Asp Leu Arg Asp Lys
      130      135      140
Lys Val Glu Tyr Val Asn Ser Met Ala Thr Phe Arg Asp Ser His Thr
      145      150      155      160
Ile Glu Tyr Val Ala Met Pro Gly Ala Glu His Arg Gln Val Thr Ser
      165      170      175
Glu Tyr Val Val Val Ala Val Gly Gly Arg Pro Arg Tyr Pro Asp Ile
      180      185      190
Pro Gly Ala Val Glu Leu Gly Ile Thr Ser Asp Asp Ile Phe Ser Tyr
      195      200      205
Glu Arg Glu Pro Gly Arg Thr Leu Val Val Gly Ala Gly Tyr Val Gly
      210      215      220
Leu Glu Cys Ala Cys Phe Leu Lys Gly Leu Gly Tyr Glu Pro Thr Val
      225      230      235      240
Met Val Arg Ser Ile Val Leu Arg Gly Phe Asp Arg Gln Met Ser Glu
      245      250      255
Leu Leu Ala Ala Met Met Thr Glu Arg Gly Ile Pro Phe Leu Gly Thr
      260      265      270
Thr Ile Pro Lys Ala Val Glu Arg Gln Ala Asp Gly Arg Leu Leu Val
      275      280      285
Arg Tyr Arg Asn Thr Thr Thr Gln Met Asp Gly Ser Asp Val Phe Asp
      290      295      300

```

Thr Val Leu Trp Ala Ile Gly Arg Lys Gly Leu Ile Glu Asp Leu Asn
 305 310 315 320
 Leu Asp Ala Ala Gly Val Lys Thr His Asp Asp Lys Ile Val Val Asp
 325 330 335
 Ala Ala Glu Ala Thr Ser Val Pro His Ile Phe Ala Val Gly Asp Ile
 340 345 350
 Ile Tyr Gly Arg Pro Glu Leu Thr Pro Val Ala Ile Leu Ser Gly Arg
 355 360 365
 Leu Leu Ala Arg Arg Leu Phe Ala Gly Ser Thr Gln Leu Met Asp Tyr
 370 375 380
 Ala Asp Val Ala Thr Thr Val Phe Thr Pro Leu Glu Tyr Ser Cys Val
 385 390 395 400
 Gly Met Ser Glu Glu Thr Ala Ile Glu Leu Arg Gly Ala Asp Asn Ile
 405 410 415
 Glu Val Phe His Gly Tyr Tyr Lys Pro Thr Glu Phe Phe Ile Pro Gln
 420 425 430
 Lys Ser Val Arg His Cys Tyr Leu Lys Ala Val Ala Glu Val Ser Gly
 435 440 445
 Asp Gln Lys Ile Leu Gly Leu His Tyr Ile Gly Pro Val Ala Gly Glu
 450 455 460
 Val Ile Gln Gly Phe Ala Ala Ala Leu Lys Thr Gly Leu Thr Val Lys
 465 470 475 480
 Thr Leu Leu Asn Thr Val Gly Ile His Pro Thr Thr Ala Glu Glu Phe
 485 490 495
 Thr Arg Leu Ser Ile Thr Lys Arg Ser Gly Arg Asp Pro Thr Pro Ala
 500 505 510
 Ser Cys Cys Ser
 515

<210> 290
 <211> 524
 <212> PRT
 <213> Homo sapiens

<220>
 <221> VARIANT
 <222> 523
 <223> Xaa = Any Amino Acid

<400> 290
 Met Ala Ala Met Ala Val Ala Leu Arg Gly Leu Gly Gly Arg Phe Arg
 1 5 10 15
 Trp Arg Thr Gln Ala Val Ala Gly Gly Val Arg Gly Ala Ala Arg Gly
 20 25 30
 Ala Ala Ala Gly Gln Arg Asp Tyr Asp Leu Leu Val Val Gly Gly Gly
 35 40 45
 Ser Gly Gly Leu Ala Cys Ala Lys Glu Ala Ala Gln Leu Gly Arg Lys
 50 55 60
 Val Ser Val Val Asp Tyr Val Glu Pro Ser Pro Gln Gly Thr Arg Trp
 65 70 75 80
 Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys Ile Pro Lys Lys Leu
 85 90 95
 Met His Gln Ala Ala Leu Leu Gly Gly Leu Ile Gln Asp Ala Pro Asn
 100 105 110
 Tyr Gly Trp Glu Val Ala Gln Pro Val Pro His Asp Trp Arg Lys Met
 115 120 125
 Ala Glu Ala Val Gln Asn His Val Lys Ser Leu Asn Trp Gly His Arg
 130 135 140
 Val Gln Leu Gln Asp Arg Lys Val Lys Tyr Phe Asn Ile Lys Ala Ser
 145 150 155 160
 Phe Val Asp Glu His Thr Val Cys Gly Val Ala Lys Gly Gly Lys Glu
 165 170 175
 Ile Leu Leu Ser Ala Asp His Ile Ile Ala Thr Gly Gly Arg Pro
 180 185 190
 Arg Tyr Pro Thr His Ile Glu Gly Ala Leu Glu Tyr Gly Ile Thr Ser
 195 200 205

Asp 210 Ile Phe Trp Leu Lys Glu Ser Pro Gly Lys Thr Leu Val Val
 215 220
 Gly 225 Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe Leu Thr Gly Ile
 230 235 240
 Gly 245 Leu Asp Thr Thr Ile Met Met Arg Ser Ile Pro Leu Arg Gly Phe
 250 255
 Asp 260 Gln Gln Met Ser Ser Met Val Ile Glu His Met Ala Ser His Gly
 265 270
 Thr 275 Arg Phe Leu Arg Gly Cys Ala Pro Ser Arg Val Arg Arg Leu Pro
 280 285
 Asp 290 Gly Gln Leu Gln Val Thr Trp Glu Asp Ser Thr Thr Gly Lys Glu
 295 300
 Asp 305 Thr Gly Thr Phe Asp Thr Val Leu Trp Ala Ile Gly Arg Val Pro
 310 315 320
 Asp 325 Thr Arg Ser Leu Asn Leu Glu Lys Ala Gly Val Asp Thr Ser Pro
 330 335
 Asp 340 Thr Gln Lys Ile Leu Val Asp Ser Arg Glu Ala Thr Ser Val Pro
 345 350
 His 355 Ile Tyr Ala Ile Gly Asp Val Val Glu Gly Arg Pro Glu Leu Thr
 360 365
 Pro 370 Ile Ala Ile Met Ala Gly Arg Leu Leu Val Gln Arg Leu Phe Gly
 375 380
 Gly 385 Ser Ser Asp Leu Met Asp Tyr Asp Asn Val Pro Thr Thr Val Phe
 390 395 400
 Thr 405 Pro Leu Glu Tyr Gly Cys Val Gly Leu Ser Glu Glu Glu Ala Val
 410 415
 Ala 420 Arg His Gly Gln Glu His Val Glu Val Tyr His Ala His Tyr Lys
 425 430
 Pro 435 Leu Glu Phe Thr Val Ala Gly Arg Asp Ala Ser Gln Cys Tyr Val
 440 445
 Lys 450 Met Val Cys Leu Arg Glu Pro Pro Gln Leu Val Leu Gly Leu His
 455 460
 Phe 465 Leu Gly Pro Asn Ala Gly Glu Val Thr Gln Gly Phe Ala Leu Gly
 470 475 480
 Ile 485 Lys Cys Gly Ala Ser Tyr Ala Gln Val Met Arg Thr Val Gly Ile
 490 495
 His 500 Pro Thr Cys Ser Glu Glu Val Val Lys Leu Arg Ile Ser Lys Arg
 505 510
 Ser 515 Gly Leu Asp Pro Thr Val Thr Gly Cys Xaa Gly
 520

<210> 291

<211> 497

<212> PRT

<213> Homo sapiens

<400> 291

Met 1 Asn Gly Pro Glu 5 Asp Leu Pro Lys 10 Ser Tyr Asp Tyr Asp 15 Leu Ile
 Ile 20 Gly Gly Gly Ser Gly Gly Leu 25 Ala Ala Ala Lys 30 Glu Ala Ala
 Gln 35 Tyr Gly Lys Lys Val Met Val 40 Leu Asp Phe Val 45 Thr Pro Thr Pro
 Leu 50 Gly Thr Arg Trp Gly 55 Leu Gly Gly Thr 60 Cys Val Asn Val Gly Cys
 Ile 65 Pro Lys Lys Leu 70 Met His Gln Ala 75 Leu Gly Gln Ala Leu
 Gln 85 Asp Ser Arg Asn 90 Tyr Gly Trp Lys Val 95 Glu Glu Thr Val Lys His
 Asp 100 Trp Asp Arg Met Ile Glu Ala Val 105 Gln Asn His Ile 110 Gly Ser Leu
 Asn 115 Trp Gly Tyr Arg Val Ala Leu 120 Arg Glu Lys Lys Val 125 Tyr Glu
 Asn 130 Ala Tyr Gly Gln Phe Ile 135 Gly Pro His Arg Ile 140 Lys Ala Thr Asn
 Asn 145 Lys Gly Lys Glu Lys Ile Tyr Ser Ala Glu Arg Phe Leu Ile Ala

```

145          150          155          160
Thr Gly Glu Arg Pro Arg Tyr Leu Gly Ile Pro Gly Asp Lys Glu Tyr
165          170          175
Cys Ile Ser Ser Asp Asp Leu Phe Ser Leu Pro Tyr Cys Pro Gly Lys
180          185          190
Thr Leu Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe
195          200          205
Leu Ala Gly Ile Gly Leu Asn Val Thr Val Met Val Arg Ser Ile Leu
210          215          220
Leu Arg Gly Phe Asp Gln Asp Met Ala Asn Lys Ile Gly Glu His Met
225          230          235
Glu Glu His Gly Ile Lys Phe Ile Arg Gln Phe Val Pro Ile Lys Val
245          250          255
Glu Gln Ile Glu Ala Gly Thr Pro Gly Arg Leu Arg Val Val Ala Gln
260          265          270
Ser Thr Asn Ser Glu Glu Ile Ile Glu Gly Glu Tyr Asn Thr Val Met
275          280          285
Leu Ala Ile Gly Arg Asp Ala Cys Thr Arg Lys Ile Gly Leu Glu Thr
290          295          300
Val Gly Val Lys Ile Asn Glu Lys Thr Gly Lys Ile Pro Val Thr Asp
305          310          315
Glu Glu Gln Thr Asn Val Pro Tyr Ile Tyr Ala Ile Gly Asp Ile Leu
325          330          335
Glu Asp Lys Val Glu Leu Thr Pro Val Ala Ile Gln Ala Gly Arg Leu
340          345          350
Leu Ala Gln Arg Leu Tyr Ala Gly Ser Thr Val Lys Cys Asp Tyr Glu
355          360          365
Asn Val Pro Thr Thr Val Phe Thr Pro Leu Glu Tyr Gly Ala Cys Gly
370          375          380
Leu Ser Glu Glu Lys Ala Val Glu Lys Phe Gly Glu Glu Asn Ile Glu
385          390          395
Val Tyr His Ser Tyr Phe Trp Pro Leu Glu Trp Thr Ile Pro Ser Arg
405          410          415
Asp Asn Asn Lys Cys Tyr Ala Lys Ile Cys Asn Thr Lys Asp Asn
420          425          430
Glu Arg Val Val Gly Phe His Val Leu Gly Pro Asn Ala Gly Glu Val
435          440          445
Thr Gln Gly Phe Ala Ala Ala Leu Lys Cys Gly Leu Thr Lys Lys Gln
450          455          460
Leu Asp Ser Thr Ile Gly Ile His Pro Val Cys Ala Glu Val Phe Thr
465          470          475
Thr Leu Ser Val Thr Lys Arg Ser Gly Ala Arg Ile Leu Gln Ala Gly
485          490          495
Cys

```

<210> 292

<211> 497

<212> PRT

<213> Homo sapien

<400> 292

```

Met Asn Gly Pro Glu Asp Leu Pro Lys Ser Tyr Asp Tyr Asp Leu Ile
1      5      10      15
Ile Ile Gly Gly Gly Ser Gly Gly Leu Ala Ala Ala Lys Glu Pro Ala
20      25      30
Gln Tyr Gly Lys Lys Val Met Val Leu Asp Phe Gly Thr Pro Thr Pro
35      40      45
Leu Gly Thr Arg Trp Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys
50      55      60
Ile Pro Lys Lys Leu Met His Gln Ala Ala Leu Leu Gly Gln Ala Leu
65      70      75
Gln Asp Ser Arg Asn Tyr Gly Trp Lys Val Glu Glu Thr Val Lys His
85      90      95
Asp Trp Asp Arg Met Ile Glu Ala Val Gln Asn His Ile Gly Ser Leu
100     105     110

```

Asn Trp Gly Tyr Arg Val Ala Leu Arg Glu Lys Lys Val Val Tyr Glu
 115 120 125
 Asn Ala Tyr Gly Gln Phe Ile Gly Pro His Arg Ile Lys Ala Thr Asn
 130 135 140
 Asn Lys Gly Lys Glu Lys Ile Tyr Ser Ala Glu Arg Phe Leu Ile Ala
 145 150 155 160
 Thr Gly Glu Arg Pro Arg Tyr Leu Gly Ile Pro Gly Asp Lys Glu Tyr
 165 170 175
 Cys Ile Ser Ser Asp Asp Leu Phe Ser Leu Pro Tyr Cys Pro Gly Lys
 180 185 190
 Thr Leu Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe
 195 200 205
 Leu Ala Gly Ile Gly Leu Asp Val Thr Val Met Val Arg Ser Ile Leu
 210 215 220
 Leu Arg Gly Phe Asp Gln Asp Met Ala Asn Lys Ile Gly Glu His Met
 225 230 235 240
 Glu Glu His Gly Ile Lys Phe Ile Arg Gln Phe Val Pro Ile Lys Val
 245 250 255
 Glu Gln Ile Glu Ala Gly Thr Pro Gly Arg Leu Arg Val Val Ala Gln
 260 265 270
 Ser Thr Asn Ser Glu Glu Ile Ile Glu Gly Glu Tyr Asn Thr Val Met
 275 280 285
 Leu Ala Ile Gly Arg Asp Ala Cys Thr Arg Lys Ile Gly Leu Glu Thr
 290 295 300
 Val Gly Val Lys Ile Asn Glu Lys Thr Gly Lys Ile Pro Val Thr Asp
 305 310 315 320
 Glu Glu Gln Thr Asn Val Pro Tyr Ile Tyr Ala Ile Gly Asp Ile Leu
 325 330 335
 Glu Asp Lys Val Glu Leu Thr Pro Val Ala Ile Gln Ala Gly Arg Leu
 340 345 350
 Leu Ala Gln Arg Leu Tyr Ala Gly Ser Thr Val Lys Cys Asp Tyr Glu
 355 360 365
 Asn Val Pro Thr Thr Val Phe Thr Pro Leu Glu Tyr Gly Ala Cys Gly
 370 375 380
 Leu Ser Glu Glu Lys Ala Val Glu Lys Phe Gly Glu Glu Asn Ile Glu
 385 390 395 400
 Val Tyr His Ser Tyr Phe Trp Pro Leu Glu Trp Thr Ile Pro Ser Arg
 405 410 415
 Asp Asn Asn Lys Cys Tyr Ala Lys Ile Cys Asn Thr Lys Asp Asn
 420 425 430
 Glu Arg Val Val Gly Phe His Val Leu Gly Pro Asn Ala Gly Glu Val
 435 440 445
 Thr Gln Gly Phe Ala Ala Ala Leu Lys Cys Gly Leu Thr Lys Lys Gln
 450 455 460
 Leu Asp Ser Thr Ile Gly Ile His Pro Val Cys Ala Glu Val Phe Thr
 465 470 475 480
 Thr Leu Ser Val Thr Lys Arg Ser Gly Ala Ser Ile Leu Gln Ala Gly
 485 490 495
 Cys

<210> 293
 <211> 521
 <212> PRT
 <213> Homo sapiens

<220>
 <221> VARIANT
 <222> 520
 <223> Xaa = Any Amino Acid

<400> 293
 Met Ala Val Ala Leu Arg Gly Leu Gly Gly Arg Phe Arg Trp Arg Thr
 1 5 10 15
 Gln Ala Val Ala Gly Gly Val Arg Gly Ala Ala Arg Gly Ala Ala Ala
 20 25 30

Gly Gln Arg Asp Tyr Asp Leu Leu Val Val Gly Gly Gly Ser Gly Gly
 35 40 45
 Leu Ala Cys Ala Lys Glu Ala Ala Gln Leu Gly Arg Lys Val Ala Val
 50 55 60
 Val Asp Tyr Val Glu Pro Ser Pro Gln Gly Thr Arg Trp Gly Leu Gly
 65 70 75 80
 Gly Thr Cys Val Asn Val Gly Cys Ile Pro Lys Lys Leu Met His Gln
 85 90 95
 Ala Ala Leu Leu Gly Gly Leu Ile Gln Asp Ala Pro Asn Tyr Gly Trp
 100 105 110
 Glu Val Ala Gln Pro Val Pro His Asp Trp Arg Lys Met Ala Glu Ala
 115 120 125
 Val Gln Asn His Val Lys Ser Leu Asn Trp Gly His Arg Val Gln Leu
 130 135 140
 Gln Asp Arg Lys Val Lys Tyr Phe Asn Ile Lys Ala Ser Phe Val Asp
 145 150 155 160
 Glu His Thr Val Cys Gly Val Ala Lys Gly Gly Lys Glu Ile Leu Leu
 165 170 175
 Ser Ala Asp His Ile Ile Ala Thr Gly Gly Arg Pro Arg Tyr Pro
 180 185 190
 Thr His Ile Glu Gly Ala Leu Glu Tyr Gly Ile Thr Ser Asp Asp Ile
 195 200 205
 Phe Trp Leu Lys Glu Ser Pro Gly Lys Thr Leu Val Gly Ala Ser
 210 215 220
 Tyr Val Ala Leu Glu Cys Ala Gly Phe Leu Thr Gly Ile Gly Leu Asp
 225 230 235 240
 Thr Thr Ile Met Met Arg Ser Ile Pro Leu Arg Gly Phe Asp Gln Gln
 245 250 255
 Met Ser Ser Met Val Ile Glu His Met Ala Ser His Gly Thr Arg Phe
 260 265 270
 Leu Arg Gly Cys Ala Pro Ser Arg Val Arg Arg Leu Pro Asp Gly Gln
 275 280 285
 Leu Gln Val Thr Trp Glu Asp Ser Thr Thr Gly Lys Glu Asp Thr Gly
 290 295 300
 Thr Phe Asp Thr Val Leu Trp Ala Ile Gly Arg Val Pro Asp Thr Arg
 305 310 315 320
 Ser Leu Asn Leu Glu Lys Ala Gly Val Asp Thr Ser Pro Asp Thr Gln
 325 330 335
 Lys Ile Leu Val Asp Ser Arg Glu Ala Thr Ser Val Pro His Ile Tyr
 340 345 350
 Ala Ile Gly Asp Val Val Glu Gly Arg Pro Glu Leu Thr Pro Ile Ala
 355 360 365
 Ile Met Ala Gly Arg Leu Leu Val Gln Arg Leu Phe Gly Gly Ser Ser
 370 375 380
 Asp Leu Met Asp Tyr Asp Asn Val Pro Thr Thr Val Phe Thr Pro Leu
 385 390 395 400
 Glu Tyr Gly Cys Val Gly Leu Ser Glu Glu Glu Ala Val Ala Arg His
 405 410 415
 Gly Gln Glu His Val Glu Val Tyr His Ala His Tyr Lys Pro Leu Glu
 420 425 430
 Phe Thr Val Ala Gly Arg Asp Ala Ser Gln Cys Tyr Val Lys Met Val
 435 440 445
 Cys Leu Arg Glu Pro Pro Gln Leu Val Leu Gly Leu His Phe Leu Gly
 450 455 460
 Pro Asn Ala Gly Glu Val Thr Gln Gly Phe Ala Leu Gly Ile Lys Cys
 465 470 475 480
 Gly Ala Ser Tyr Ala Gln Val Met Arg Thr Val Gly Ile His Pro Thr
 485 490 495
 Cys Ser Glu Glu Val Val Lys Leu Arg Ile Ser Lys Arg Ser Gly Leu
 500 505 510
 Asp Pro Thr Val Thr Gly Cys Xaa Gly
 515 520

<210> 294

<211> 579

<212> PRT

<213> Homo sapiens

<220>

<221> VARIANT

<222> 578

<223> Xaa = Any Amino Acid

<400> 294

Ala	Glu	Arg	Val	Val	Ile	Phe	Ser	Lys	Ser	Tyr	Cys	Pro	His	Ser	Thr
1				5					10					15	
Arg	Val	Lys	Glu	Leu	Phe	Ser	Ser	Leu	Gly	Val	Glu	Cys	Asn	Val	Leu
			20					25					30		
Glu	Leu	Asp	Gln	Val	Asp	Asp	Gly	Ala	Arg	Val	Gln	Glu	Val	Leu	Ser
		35					40					45			
Glu	Ile	Thr	Asn	Gln	Lys	Thr	Val	Pro	Asn	Ile	Phe	Val	Asn	Lys	Val
	50					55					60				
His	Val	Gly	Gly	Cys	Asp	Gln	Thr	Phe	Gln	Ala	Tyr	Gln	Ser	Gly	Leu
	65				70					75				80	
Leu	Gln	Lys	Leu	Leu	Gln	Glu	Asp	Leu	Ala	Tyr	Asp	Tyr	Asp	Leu	Ile
			85						90					95	
Ile	Ile	Gly	Gly	Gly	Ser	Gly	Gly	Leu	Ser	Cys	Ala	Lys	Glu	Ala	Ala
		100						105					110		
Ile	Leu	Gly	Lys	Lys	Val	Met	Val	Leu	Asp	Phe	Val	Val	Pro	Ser	Pro
	115						120					125			
Gln	Gly	Thr	Ser	Trp	Gly	Leu	Gly	Gly	Thr	Cys	Val	Asn	Val	Gly	Cys
	130					135					140				
Ile	Pro	Lys	Lys	Leu	Met	His	Gln	Ala	Ala	Leu	Leu	Gly	Gln	Ala	Leu
	145				150					155				160	
Cys	Asp	Ser	Arg	Lys	Phe	Gly	Trp	Glu	Tyr	Asn	Gln	Gln	Val	Arg	His
			165						170					175	
Asn	Trp	Glu	Thr	Met	Thr	Lys	Ala	Ile	Gln	Asn	His	Ile	Ser	Ser	Leu
	180						185					190			
Asn	Trp	Gly	Tyr	Arg	Leu	Ser	Leu	Arg	Glu	Lys	Ala	Val	Ala	Tyr	Val
	195						200					205			
Asn	Ser	Tyr	Gly	Glu	Phe	Val	Glu	His	His	Lys	Ile	Lys	Ala	Thr	Asn
	210					215					220				
Lys	Lys	Gly	Gln	Glu	Thr	Tyr	Tyr	Thr	Ala	Ala	Gln	Phe	Val	Ile	Ala
	225				230					235				240	
Thr	Gly	Glu	Arg	Pro	Arg	Tyr	Leu	Gly	Ile	Gln	Gly	Asp	Lys	Glu	Tyr
			245						250					255	
Cys	Ile	Thr	Ser	Asp	Asp	Leu	Phe	Ser	Leu	Pro	Tyr	Cys	Pro	Gly	Lys
		260					265					270			
Thr	Leu	Val	Val	Gly	Ala	Ser	Tyr	Val	Ala	Leu	Glu	Cys	Ala	Gly	Phe
	275						280					285			
Leu	Ala	Gly	Phe	Gly	Leu	Asp	Val	Thr	Val	Met	Val	Arg	Ser	Ile	Leu
	290					295					300				
Leu	Arg	Gly	Phe	Asp	Gln	Glu	Met	Ala	Glu	Lys	Val	Gly	Ser	Tyr	Met
	305				310					315				320	
Glu	Gln	His	Gly	Val	Lys	Phe	Leu	Arg	Lys	Phe	Ile	Pro	Val	Met	Val
			325						330					335	
Gln	Gln	Leu	Glu	Lys	Gly	Ser	Pro	Gly	Lys	Leu	Lys	Val	Leu	Ala	Lys
		340						345					350		
Ser	Thr	Glu	Gly	Thr	Glu	Thr	Ile	Glu	Gly	Val	Tyr	Asn	Thr	Val	Leu
	355						360					365			
Leu	Ala	Ile	Gly	Arg	Asp	Ser	Cys	Thr	Arg	Lys	Ile	Gly	Leu	Glu	Lys
	370					375					380				
Ile	Gly	Val	Lys	Ile	Asn	Glu	Lys	Ser	Gly	Lys	Ile	Pro	Val	Asn	Asp
	385				390					395				400	
Val	Glu	Gln	Thr	Asn	Val	Pro	Tyr	Val	Tyr	Ala	Val	Gly	Asp	Ile	Leu
			405						410					415	
Glu	Asp	Lys	Pro	Glu	Leu	Thr	Pro	Val	Ala	Ile	Gln	Ser	Gly	Lys	Leu
		420						425					430		
Leu	Ala	Gln	Arg	Leu	Phe	Gly	Ala	Ser	Leu	Glu	Lys	Cys	Asp	Tyr	Ile
	435						440					445			
Asn	Val	Pro	Thr	Thr	Val	Phe	Thr	Pro	Leu	Glu	Tyr	Gly	Cys	Cys	Gly
	450					455					460				
Leu	Ser	Glu	Glu	Lys	Ala	Ile	Glu	Val	Tyr	Lys	Lys	Glu	Asn	Leu	Glu

[illegible]

```
<210> 295
<211> 524
<212> PRT
<213> Homo sapien
```

```
<220>  
<221> VARIANT  
<222> 523  
<223> Xaa = Any Amino Acid
```

<400>	295														
Met	Ala	Ala	Met	Ala	Val	Ala	Leu	Arg	Gly	Leu	Gly	Gly	Arg	Phe	Arg
1				5					10					15	
Trp	Arg	Thr	Gln	Ala	Val	Ala	Gly	Gly	Val	Arg	Gly	Ala	Ala	Arg	Gly
			20					25					30		
Ala	Ala	Ala	Gly	Gln	Arg	Asp	Tyr	Asp	Leu	Leu	Val	Val	Gly	Gly	Gly
		35					40					45			
Ser	Gly	Gly	Leu	Ala	Cys	Ala	Lys	Glu	Ala	Ala	Gln	Leu	Gly	Arg	Lys
	50					55					60				
Val	Ala	Val	Val	Asp	Tyr	Val	Glu	Pro	Ser	Pro	Gln	Gly	Thr	Arg	Trp
65					70					75					80
Gly	Leu	Gly	Gly	Thr	Cys	Val	Asn	Val	Gly	Cys	Ile	Pro	Lys	Lys	Leu
				85					90					95	
Met	His	Gln	Ala	Ala	Leu	Leu	Gly	Gly	Leu	Ile	Gln	Asp	Ala	Pro	Asn
			100					105					110		
Tyr	Gly	Trp	Glu	Val	Ala	Gln	Pro	Val	Pro	His	Asp	Trp	Arg	Lys	Met
		115					120					125			
Ala	Glu	Ala	Val	Gln	Asn	His	Val	Lys	Ser	Leu	Asn	Trp	Gly	His	Arg
	130					135					140				
Val	Gln	Leu	Gln	Asp	Arg	Lys	Val	Lys	Tyr	Phe	Asn	Ile	Lys	Ala	Ser
145					150					155					160
Phe	Val	Asp	Glu	His	Thr	Val	Cys	Gly	Val	Ala	Lys	Gly	Gly	Lys	Glu
				165					170					175	
Ile	Leu	Leu	Ser	Ala	Asp	His	Ile	Ile	Ile	Ala	Thr	Gly	Gly	Arg	Pro
			180					185					190		
Arg	Tyr	Pro	Thr	His	Ile	Glu	Gly	Ala	Leu	Glu	Tyr	Gly	Ile	Thr	Ser
		195					200					205			
Asp	Asp	Ile	Phe	Trp	Leu	Lys	Glu	Ser	Pro	Gly	Lys	Thr	Leu	Val	Val
	210					215					220				
Gly	Ala	Ser	Tyr	Val	Ala	Leu	Glu	Cys	Ala	Gly	Phe	Leu	Thr	Gly	Ile
225					230					235					240
Gly	Leu	Asp	Thr	Thr	Ile	Met	Met	Arg	Ser	Ile	Pro	Leu	Arg	Gly	Phe
				245					250					255	
Asp	Gln	Gln	Met	Ser	Ser	Met	Val	Ile	Glu	His	Met	Ala	Ser	His	Gly
			260					265					270		
Thr	Arg	Phe	Leu	Arg	Gly	Cys	Ala	Pro	Ser	Arg	Val	Arg	Arg	Leu	Pro
		275					280					285			
Asp	Gly	Gln	Leu	Gln	Val	Thr	Trp	Glu	Asp	Ser	Thr	Thr	Gly	Lys	Glu
	290					295					300				
Asp	Thr	Gly	Thr	Phe	Asp	Thr	Val	Leu	Trp	Ala	Ile	Gly	Arg	Val	Pro

```

305          310          315          320
Asp Thr Arg Ser Leu Asn Leu Glu Lys Ala Gly Val Asp Thr Ser Pro
          325          330          335
Asp Thr Gln Lys Ile Leu Val Asp Ser Arg Glu Ala Thr Ser Val Pro
          340          345          350
His Ile Tyr Ala Ile Gly Asp Val Val Glu Gly Arg Pro Glu Leu Thr
          355          360          365
Pro Ile Ala Ile Met Ala Gly Arg Leu Leu Val Gln Arg Leu Phe Gly
          370          375          380
Gly Ser Ser Asp Leu Met Asp Tyr Asp Asn Val Pro Thr Thr Val Phe
385          390          395          400
Thr Pro Leu Glu Tyr Gly Cys Val Gly Leu Ser Glu Glu Glu Ala Val
          405          410          415
Ala Arg His Gly Gln Glu His Val Glu Val Tyr His Ala His Tyr Lys
          420          425          430
Pro Leu Glu Phe Thr Val Ala Gly Arg Asp Ala Ser Gln Cys Tyr Val
          435          440          445
Lys Met Val Cys Leu Arg Glu Pro Pro Gln Leu Val Leu Gly Leu His
          450          455          460
Phe Leu Gly Pro Asn Ala Gly Glu Val Thr Gln Gly Phe Ala Leu Gly
465          470          475          480
Ile Lys Cys Gly Ala Ser Tyr Ala Gln Val Met Arg Thr Val Gly Ile
          485          490          495
His Pro Thr Cys Ser Glu Glu Val Val Lys Leu Arg Ile Ser Lys Arg
          500          505          510
Ser Gly Leu Asp Pro Thr Val Thr Gly Cys Xaa Gly
          515          520

```

<210> 296
 <211> 577
 <212> PRT
 <213> Homo sapien

<220>
 <221> VARIANT
 <222> 576
 <223> Xaa = Any Amino Acid

```

<400> 296
Arg Val Val Ile Phe Ser Lys Ser Tyr Cys Pro His Ser Thr Arg Val
1          5          10          15
Lys Glu Leu Phe Ser Ser Leu Gly Val Glu Cys Asn Val Leu Glu Leu
          20          25          30
Asp Gln Val Asp Asp Gly Ala Arg Val Gln Glu Val Leu Ser Glu Ile
          35          40          45
Thr Asn Gln Lys Thr Val Pro Asn Ile Phe Val Asn Lys Val His Val
          50          55          60
Gly Gly Cys Asp Gln Thr Phe Gln Ala Tyr Gln Ser Gly Leu Leu Gln
65          70          75          80
Lys Leu Leu Gln Glu Asp Leu Ala Tyr Asp Tyr Asp Leu Ile Ile Ile
          85          90          95
Gly Gly Gly Ser Gly Gly Leu Ser Cys Ala Lys Glu Ala Ala Ile Leu
          100          105          110
Gly Lys Lys Val Met Val Leu Asp Phe Val Val Pro Ser Pro Gln Gly
          115          120          125
Thr Ser Trp Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys Ile Pro
          130          135          140
Lys Lys Leu Met His Gln Ala Ala Leu Leu Gly Gln Ala Leu Cys Asp
145          150          155          160
Ser Arg Lys Phe Gly Trp Glu Tyr Asn Gln Gln Val Arg His Asn Trp
          165          170          175
Glu Thr Met Thr Lys Ala Ile Gln Asn His Ile Ser Ser Leu Asn Trp
          180          185          190
Gly Tyr Arg Leu Ser Leu Arg Glu Lys Ala Val Ala Tyr Val Asn Ser
          195          200          205
Tyr Gly Glu Phe Val Glu His His Lys Ile Lys Ala Thr Asn Lys Lys

```

```

      210      215      220
Gly Gln Glu Thr Tyr Tyr Thr Ala Ala Gln Phe Val Ile Ala Thr Gly
225      230      235      240
Glu Arg Pro Arg Tyr Leu Gly Ile Gln Gly Asp Lys Glu Tyr Cys Ile
      245      250      255
Thr Ser Asp Asp Leu Phe Ser Leu Pro Tyr Cys Pro Gly Lys Pro Leu
      260      265      270
Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe Leu Ala
      275      280      285
Gly Phe Gly Leu Asp Val Thr Val Met Val Arg Ser Ile Leu Leu Arg
      290      295      300
Gly Phe Asp Gln Glu Met Ala Glu Lys Val Gly Ser Tyr Met Glu Gln
305      310      315      320
His Gly Val Lys Phe Leu Arg Lys Phe Ile Pro Val Met Val Gln Gln
      325      330      335
Leu Glu Lys Gly Ser Pro Gly Lys Leu Lys Val Leu Ala Lys Ser Thr
      340      345      350
Glu Gly Thr Glu Thr Ile Glu Gly Val Tyr Asn Thr Val Leu Leu Ala
      355      360      365
Ile Gly Arg Asp Ser Cys Thr Arg Lys Ile Gly Leu Glu Lys Ile Gly
      370      375      380
Val Lys Ile Asn Glu Lys Ser Gly Lys Ile Pro Val Asn Asp Val Glu
385      390      395      400
Gln Thr Asn Val Pro Tyr Val Tyr Ala Val Gly Asp Ile Leu Glu Asp
      405      410      415
Lys Pro Glu Leu Thr Pro Val Ala Ile Gln Ser Gly Lys Leu Leu Ala
      420      425      430
Gln Arg Leu Phe Gly Ala Ser Leu Glu Lys Cys Asp Tyr Ile Asn Val
      435      440      445
Pro Thr Thr Val Phe Thr Pro Leu Glu Tyr Gly Cys Cys Gly Leu Ser
      450      455      460
Glu Glu Lys Ala Ile Glu Val Tyr Lys Lys Glu Asn Leu Glu Ile Tyr
465      470      475      480
His Thr Leu Phe Trp Pro Leu Glu Trp Thr Val Ala Gly Arg Glu Asn
      485      490      495
Asn Thr Cys Tyr Ala Lys Ile Ile Cys Asn Lys Phe Asp His Asp Arg
      500      505      510
Val Ile Gly Phe His Ile Leu Gly Pro Asn Ala Gly Glu Val Thr Gln
      515      520      525
Gly Phe Ala Ala Ala Met Lys Cys Gly Leu Thr Lys Gln Leu Leu Asp
      530      535      540
Asp Thr Ile Gly Ile His Pro Thr Cys Gly Glu Val Phe Thr Thr Leu
545      550      555      560
Glu Ile Thr Lys Ser Ser Gly Leu Asp Ile Thr Gln Lys Gly Cys Xaa
      565      570      575
Gly

```

<210> 297

<211> 494

<212> PRT

<213> Homo sapien

<400> 297

```

Met Glu Asp Gln Ala Gly Gln Arg Asp Tyr Asp Leu Leu Val Val Gly
1      5      10      15
Gly Gly Ser Gly Leu Ala Cys Ala Lys Glu Ala Ala Gln Leu Gly
      20      25      30
Arg Lys Val Ala Val Val Asp Tyr Val Glu Pro Ser Pro Gln Gly Thr
      35      40      45
Arg Trp Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys Ile Pro Lys
50      55      60
Lys Leu Met His Gln Ala Ala Leu Leu Gly Gly Leu Ile Gln Asp Ala
65      70      75      80
Pro Asn Tyr Gly Trp Glu Val Ala Gln Pro Val Pro His Asp Trp Arg
      85      90      95

```

Lys Met Ala Glu Ala Val Gln Asn His Val Lys Ser Leu Asn Trp Gly
 100 105 110
 His Arg Val Gln Leu Gln Asp Arg Lys Val Lys Tyr Phe Asn Ile Lys
 115 120 125
 Ala Ser Phe Val Asp Glu His Thr Val Cys Gly Val Ala Lys Gly Gly
 130 135 140
 Lys Glu Ile Leu Leu Ser Ala Asp His Ile Ile Ala Thr Gly Gly
 145 150 155 160
 Arg Pro Arg Tyr Pro Thr His Ile Glu Gly Ala Leu Glu Tyr Gly Ile
 165 170 175
 Thr Ser Asp Asp Ile Phe Trp Leu Lys Glu Ser Pro Gly Lys Thr Leu
 180 185 190
 Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe Leu Thr
 195 200 205
 Gly Ile Gly Leu Asp Thr Thr Ile Met Met Arg Ser Ile Pro Leu Arg
 210 215 220
 Gly Phe Asp Gln Gln Met Ser Ser Met Val Ile Glu His Met Ala Ser
 225 230 235 240
 His Gly Thr Arg Phe Leu Arg Gly Cys Ala Pro Ser Arg Val Arg Arg
 245 250 255
 Leu Pro Asp Gly Gln Leu Gln Val Thr Trp Glu Asp Ser Thr Thr Gly
 260 265 270
 Lys Glu Asp Thr Gly Thr Phe Asp Thr Val Leu Trp Ala Ile Gly Arg
 275 280 285
 Val Pro Asp Thr Arg Ser Leu Asn Leu Glu Lys Ala Gly Val Asp Thr
 290 295 300
 Ser Pro Asp Thr Gln Lys Ile Leu Val Asp Ser Arg Glu Ala Thr Ser
 305 310 315 320
 Val Pro His Ile Tyr Ala Ile Gly Asp Val Val Glu Gly Arg Pro Glu
 325 330 335
 Leu Thr Pro Thr Ala Ile Met Ala Gly Arg Leu Leu Val Gln Arg Leu
 340 345 350
 Phe Gly Gly Ser Ser Asp Leu Met Asp Tyr Asp Asn Val Pro Thr Thr
 355 360 365
 Val Phe Thr Pro Leu Glu Tyr Gly Cys Val Gly Leu Ser Glu Glu Glu
 370 375 380
 Ala Val Ala Arg His Gly Gln Glu His Val Glu Val Tyr His Ala His
 385 390 395 400
 Tyr Lys Pro Leu Glu Phe Thr Val Ala Gly Arg Asp Ala Ser Gln Cys
 405 410 415
 Tyr Val Lys Met Val Cys Leu Arg Glu Pro Pro Gln Leu Val Leu Gly
 420 425 430
 Leu His Phe Leu Gly Pro Asn Ala Gly Glu Val Thr Gln Gly Phe Ala
 435 440 445
 Leu Gly Ile Lys Cys Gly Ala Ser Tyr Ala Gln Val Met Arg Thr Val
 450 455 460
 Gly Ile His Pro Thr Cys Ser Glu Glu Val Val Lys Leu Arg Ile Ser
 465 470 475 480
 Lys Arg Ser Gly Leu Asp Pro Thr Val Thr Gly Cys Cys Gly
 485 490

<210> 298

<211> 521

<212> PRT

<213> Homo sapien

<400> 298

Met Ala Ala Met Ala Val Ala Leu Arg Gly Leu Gly Gly Arg Phe Arg
 1 5 10 15
 Trp Arg Thr Gln Ala Val Ala Gly Gly Val Arg Gly Ala Ala Arg Gly
 20 25 30
 Ala Ala Gly Gln Arg Asp Tyr Asp Leu Leu Val Val Gly Gly Ser
 35 40 45
 Gly Gly Leu Ala Cys Ala Lys Glu Ala Ala Gln Leu Gly Arg Lys Val
 50 55 60
 Ser Val Val Asp Tyr Val Glu Pro Ser Pro Gln Gly Thr Arg Trp Gly

65	70	75	80
Leu Gly Gly Thr Cys Val Asn Val Gly Cys Ile Pro Lys Lys Leu Met			
	85	90	95
His Gln Ala Ala Leu Leu Gly Gly Leu Ile Gln Asp Ala Pro Asn Tyr			
	100	105	110
Gly Trp Glu Val Ala Gln Pro Val Pro His Asp Trp Arg Lys Met Ala			
	115	120	125
Glu Ala Val Gln Asn His Val Lys Ser Leu Asn Trp Gly His Arg Val			
	130	135	140
Gln Leu Gln Asp Arg Lys Val Lys Tyr Phe Asn Ile Lys Ala Ser Phe			
	145	150	155
Val Asp Glu His Thr Val Cys Gly Val Ala Lys Gly Gly Lys Glu Ile			
	165	170	175
Leu Leu Ser Ala Asp His Ile Ile Ile Ala Thr Gly Gly Arg Pro Arg			
	180	185	190
Tyr Pro Thr His Ile Glu Gly Ala Leu Glu Tyr Gly Ile Thr Ser Asp			
	195	200	205
Asp Ile Phe Trp Leu Lys Glu Ser Pro Gly Lys Thr Leu Val Val Gly			
	210	215	220
Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe Leu Thr Gly Ile Gly			
	225	230	235
Leu Asp Thr Thr Ile Met Met Arg Ser Ile Pro Leu Arg Gly Phe Asp			
	245	250	255
Gln Gln Met Ser Ser Met Val Ile Glu His Met Ala Ser His Gly Thr			
	260	265	270
Arg Phe Leu Arg Gly Cys Ala Pro Ser Arg Val Lys Arg Leu Pro Asp			
	275	280	285
Gly Gln Leu Gln Val Thr Trp Glu Asp Ser Thr Thr Gly Lys Glu Asp			
	290	295	300
Thr Gly Thr Phe Asp Thr Val Leu Trp Ala Ile Gly Arg Val Pro Asp			
	305	310	315
Thr Arg Ser Leu Asn Leu Glu Lys Ala Gly Val Asp Thr Ser Pro Asp			
	325	330	335
Thr Gln Lys Ile Leu Val Asp Ser Arg Glu Ala Thr Ser Val Pro His			
	340	345	350
Ile Tyr Ala Ile Gly Asp Val Val Glu Gly Arg Pro Glu Leu Thr Pro			
	355	360	365
Thr Ala Ile Met Ala Gly Arg Leu Leu Val Gln Arg Leu Phe Gly Gly			
	370	375	380
Ser Ser Asp Leu Met Asp Tyr Asp Asn Val Pro Thr Thr Val Phe Thr			
	385	390	395
Pro Leu Glu Tyr Gly Cys Val Gly Leu Ser Glu Glu Glu Ala Val Ala			
	405	410	415
Arg His Gly Gln Glu His Val Glu Val Tyr His Ala His Tyr Lys Pro			
	420	425	430
Leu Glu Phe Thr Val Ala Gly Arg Asp Ala Ser Gln Cys Tyr Val Lys			
	435	440	445
Met Val Cys Leu Arg Glu Pro Pro Gln Leu Val Leu Gly Leu His Phe			
	450	455	460
Leu Gly Pro Asn Ala Gly Glu Val Thr Gln Gly Phe Ala Leu Gly Ile			
	465	470	475
Lys Cys Gly Ala Ser Tyr Ala Gln Val Met Arg Thr Val Gly Ile His			
	485	490	495
Pro Thr Cys Ser Glu Glu Val Val Lys Leu Arg Ile Ser Lys Arg Ser			
	500	505	510
Gly Leu Asp Pro Thr Val Thr Gly Cys			
	515	520	

<210> 299

<211> 549

<212> PRT

<213> Homo sapien

<400> 299

Met	Ser	Cys	Glu	Asp	Gly	Arg	Ala	Leu	Glu	Gly	Thr	Leu	Ser	Glu	Leu
1									10					15	

Ala	Ala	Glu	Thr	Asp	Leu	Pro	Val	Val	Phe	Val	Lys	Gln	Arg	Lys	Ile
			20					25					30		
Gly	Gly	His	Gly	Pro	Thr	Leu	Lys	Ala	Tyr	Gln	Glu	Gly	Arg	Leu	Gln
		35					40					45			
Lys	Leu	Leu	Lys	Met	Asn	Gly	Pro	Glu	Asp	Leu	Pro	Lys	Ser	Tyr	Asp
		50				55					60				
Tyr	Asp	Leu	Ile	Ile	Ile	Gly	Gly	Gly	Ser	Gly	Gly	Leu	Ala	Ala	Ala
65					70				75						80
Lys	Glu	Ala	Ala	Gln	Tyr	Gly	Lys	Lys	Val	Met	Val	Leu	Asp	Phe	Val
				85					90					95	
Thr	Pro	Thr	Pro	Leu	Gly	Thr	Arg	Trp	Gly	Leu	Gly	Gly	Thr	Cys	Val
			100					105					110		
Asn	Val	Gly	Cys	Ile	Pro	Lys	Lys	Leu	Met	His	Gln	Ala	Ala	Leu	Leu
		115					120					125			
Gly	Gln	Ala	Leu	Gln	Asp	Ser	Arg	Asn	Tyr	Gly	Trp	Lys	Val	Glu	Glu
		130				135					140				
Thr	Val	Lys	His	Asp	Trp	Asp	Arg	Met	Ile	Glu	Ala	Val	Gln	Asn	His
145					150					155					160
Ile	Gly	Ser	Leu	Asn	Trp	Gly	Tyr	Arg	Val	Ala	Leu	Arg	Glu	Lys	Lys
				165					170					175	
Val	Val	Tyr	Glu	Asn	Ala	Tyr	Gly	Gln	Phe	Ile	Gly	Pro	His	Arg	Ile
			180					185					190		
Lys	Ala	Thr	Asn	Asn	Lys	Gly	Lys	Glu	Lys	Ile	Tyr	Ser	Ala	Glu	Arg
		195					200					205			
Phe	Leu	Ile	Ala	Thr	Gly	Glu	Arg	Pro	Arg	Tyr	Leu	Gly	Ile	Pro	Gly
		210				215					220				
Asp	Lys	Glu	Tyr	Cys	Ile	Ser	Ser	Asp	Asp	Leu	Phe	Ser	Leu	Pro	Tyr
225					230					235					240
Cys	Pro	Gly	Lys	Thr	Leu	Val	Val	Gly	Ala	Ser	Tyr	Val	Ala	Leu	Glu
				245					250					255	
Cys	Ala	Gly	Phe	Leu	Ala	Gly	Ile	Gly	Leu	Asp	Val	Thr	Val	Met	Val
			260					265					270		
Arg	Ser	Ile	Leu	Leu	Arg	Gly	Phe	Asp	Gln	Asp	Met	Ala	Asn	Lys	Ile
		275					280					285			
Gly	Glu	His	Met	Glu	Glu	His	Gly	Ile	Lys	Phe	Ile	Arg	Gln	Phe	Val
		290				295					300				
Pro	Ile	Lys	Val	Glu	Gln	Ile	Glu	Ala	Gly	Thr	Pro	Gly	Arg	Leu	Arg
305					310					315					320
Val	Val	Ala	Gln	Ser	Thr	Asn	Ser	Glu	Glu	Ile	Ile	Glu	Gly	Glu	Tyr
				325					330					335	
Asn	Thr	Val	Met	Leu	Ala	Ile	Gly	Arg	Asp	Ala	Cys	Thr	Arg	Lys	Ile
			340					345					350		
Gly	Leu	Glu	Thr	Val	Gly	Val	Lys	Ile	Asn	Glu	Lys	Thr	Gly	Lys	Ile
		355					360					365			
Pro	Val	Thr	Asp	Glu	Glu	Gln	Thr	Asn	Val	Pro	Tyr	Ile	Tyr	Ala	Ile
		370				375					380				
Gly	Asp	Ile	Leu	Glu	Asp	Lys	Val	Glu	Leu	Thr	Pro	Val	Ala	Ile	Gln
385					390					395					400
Ala	Gly	Arg	Leu	Leu	Ala	Gln	Arg	Leu	Tyr	Ala	Gly	Ser	Thr	Val	Lys
				405					410					415	
Cys	Asp	Tyr	Glu	Asn	Val	Pro	Thr	Thr	Val	Phe	Thr	Pro	Leu	Glu	Tyr
			420					425					430		
Gly	Ala	Cys	Gly	Leu	Ser	Glu	Glu	Lys	Ala	Val	Glu	Lys	Phe	Gly	Glu
		435					440					445			
Glu	Asn	Ile	Glu	Val	Tyr	His	Ser	Tyr	Phe	Trp	Pro	Leu	Glu	Trp	Thr
		450				455					460				
Ile	Pro	Ser	Arg	Asp	Asn	Asn	Lys	Cys	Tyr	Ala	Lys	Ile	Ile	Cys	Asn
465					470					475					480
Thr	Lys	Asp	Asn	Glu	Arg	Val	Val	Gly	Phe	His	Val	Leu	Gly	Pro	Asn
				485					490					495	
Ala	Gly	Glu	Val	Thr	Gln	Gly	Phe	Ala	Ala	Ala	Leu	Lys	Cys	Gly	Leu
			500					505					510		
Thr	Lys	Lys	Gln	Leu	Asp	Ser	Thr	Ile	Gly	Ile	His	Pro	Val	Cys	Ala
		515					520					525			
Glu	Val	Phe	Thr	Thr	Leu	Ser	Val	Thr	Lys	Arg	Ser	Gly	Ala	Ser	Ile
		530				535					540				
Leu	Gln	Ala	Gly	Cys											

545

<210> 300
 <211> 613
 <212> PRT
 <213> Mus musculus

<220>
 <221> VARIANT
 <222> 612
 <223> Xaa = Any Amino Acid

<400> 300
 Met Pro Val Asp Asp Cys Trp Leu Tyr Phe Pro Ala Ser Arg Gly Arg
 1 5 10 15
 Thr Phe Val Gln Thr Val Trp Val Ala Pro Thr Cys Pro Asn Cys Cys
 20 25 30
 Trp Phe Pro Gly Phe Leu Pro Pro Val Pro Arg Pro Pro His Val Pro
 35 40 45
 Arg Val Leu Leu Arg Gly Pro Arg Gly Ala Val Leu Pro Ala Ser Arg
 50 55 60
 Pro Ser Lys Thr Leu Pro Ser Ser Ser Gln Thr Pro Cys Pro Thr Asp
 65 70 75 80
 Pro Cys Ile Cys Pro Pro Pro Ser Thr Pro Asp Ser Arg Gln Glu Lys
 85 90 95
 Asn Thr Gln Ser Glu Leu Pro Asn Lys Lys Gly Gln Leu Gln Lys Leu
 100 105 110
 Pro Thr Met Asn Gly Ser Lys Asp Pro Pro Gly Ser Tyr Asp Phe Asp
 115 120 125
 Leu Ile Ile Ile Gly Gly Gly Ser Gly Gly Leu Ala Ala Ala Lys Glu
 130 135 140
 Ala Ala Lys Phe Asp Lys Lys Val Leu Val Leu Asp Phe Val Thr Pro
 145 150 155 160
 Thr Pro Leu Gly Thr Arg Trp Gly Leu Gly Gly Thr Cys Val Asn Val
 165 170 175
 Gly Cys Ile Pro Lys Lys Leu Met His Gln Ala Ala Leu Leu Gly Gln
 180 185 190
 Ala Leu Lys Asp Ser Arg Asn Tyr Gly Trp Lys Val Glu Asp Thr Val
 195 200 205
 Lys His Asp Trp Glu Lys Met Thr Glu Ser Val Gln Ser His Ile Gly
 210 215 220
 Ser Leu Asn Trp Gly Tyr Arg Val Ala Leu Arg Glu Lys Lys Val Val
 225 230 235 240
 Tyr Glu Asn Ala Tyr Gly Arg Phe Ile Gly Pro His Arg Ile Val Ala
 245 250 255
 Thr Asn Asn Lys Gly Lys Glu Lys Ile Tyr Ser Ala Glu Arg Phe Leu
 260 265 270
 Ile Ala Thr Gly Glu Arg Pro Arg Tyr Leu Gly Ile Pro Gly Asp Lys
 275 280 285
 Glu Tyr Cys Ile Ser Ser Asp Asp Leu Phe Ser Leu Pro Tyr Cys Pro
 290 295 300
 Gly Lys Thr Leu Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala
 305 310 315 320
 Gly Phe Leu Ala Gly Ile Gly Leu Asp Val Thr Val Met Val Arg Ser
 325 330 335
 Ile Leu Leu Arg Gly Phe Asp Gln Asp Met Ala Asn Lys Ile Gly Glu
 340 345 350
 His Met Glu Glu His Gly Ile Lys Phe Ile Arg Gln Phe Val Pro Thr
 355 360 365
 Lys Ile Glu Gln Ile Glu Ala Gly Thr Pro Gly Arg Leu Arg Val Thr
 370 375 380
 Ala Gln Ser Thr Asn Ser Glu Glu Thr Ile Glu Gly Glu Phe Asn Thr
 385 390 395 400
 Val Leu Leu Ala Val Gly Arg Asp Ser Cys Thr Arg Thr Ile Gly Leu
 405 410 415
 Glu Thr Val Gly Val Lys Ile Asn Glu Lys Thr Gly Lys Ile Pro Val

Thr	Asp	Glu	420	Gln	Thr	Asn	Val	425	Pro	Tyr	Ile	Tyr	Ala	430	Ile	Gly	Asp
		435					440						445				
Ile	Leu	Glu	Gly	Lys	Leu	Glu	Leu	Thr	Pro	Val	Ala	Ile	Gln	Ala	Gly		
	450					455					460						
Arg	Leu	Leu	Ala	Gln	Arg	Leu	Tyr	Gly	Gly	Ser	Asn	Val	Lys	Cys	Asp		
465					470					475					480		
Tyr	Asp	Asn	Val	Pro	Thr	Thr	Val	Phe	Thr	Pro	Leu	Glu	Tyr	Gly	Cys		
			485						490					495			
Cys	Gly	Leu	Ser	Glu	Glu	Lys	Ala	Val	Glu	Lys	Phe	Gly	Glu	Glu	Asn		
		500						505					510				
Ile	Glu	Val	Tyr	His	Ser	Phe	Phe	Trp	Pro	Leu	Glu	Trp	Thr	Val	Pro		
	515						520					525					
Ser	Arg	Asp	Asn	Asn	Lys	Cys	Tyr	Ala	Lys	Ile	Ile	Cys	Asn	Leu	Lys		
	530					535					540						
Asp	Asp	Glu	Arg	Val	Val	Gly	Phe	His	Val	Leu	Gly	Pro	Asn	Ala	Gly		
545					550					555					560		
Glu	Val	Thr	Gln	Gly	Phe	Ala	Ala	Ala	Leu	Lys	Cys	Gly	Leu	Thr	Lys		
			565					570					575				
Gln	Gln	Leu	Asp	Ser	Thr	Ile	Gly	Ile	His	Pro	Val	Cys	Ala	Glu	Ile		
		580						585					590				
Phe	Thr	Thr	Leu	Ser	Val	Thr	Lys	Arg	Ser	Gly	Gly	Asp	Ile	Leu	Gln		
	595						600					605					
Ser	Gly	Cys	Xaa	Gly													
	610																

<210> 301

<211> 310

<212> PRT

<213> Mus musculus

<400> 301

Met	Asn	Gly	Ser	Lys	Asp	Pro	Pro	Gly	Ser	Tyr	Asp	Phe	Asp	Leu	Ile		
1				5				10					15				
Ile	Ile	Gly	Gly	Ser	Gly	Gly	Leu	Ala	Ala	Ala	Lys	Glu	Ala	Ala			
		20					25					30					
Lys	Phe	Asp	Lys	Lys	Val	Leu	Val	Leu	Asp	Phe	Val	Thr	Pro	Thr	Pro		
	35					40					45						
Leu	Gly	Thr	Arg	Trp	Gly	Leu	Gly	Gly	Thr	Cys	Val	Asn	Val	Gly	Cys		
	50				55					60							
Ile	Pro	Lys	Lys	Leu	Met	His	Gln	Ala	Ala	Leu	Leu	Gly	Gln	Ala	Leu		
65				70						75					80		
Lys	Asp	Ser	Arg	Asn	Tyr	Gly	Trp	Lys	Val	Glu	Asp	Thr	Val	Lys	His		
			85					90						95			
Asp	Trp	Glu	Lys	Met	Thr	Glu	Ser	Val	Gln	Ser	His	Ile	Gly	Ser	Leu		
		100						105					110				
Asn	Trp	Gly	Tyr	Arg	Val	Ala	Leu	Arg	Glu	Lys	Lys	Val	Val	Tyr	Glu		
	115						120					125					
Asn	Ala	Tyr	Gly	Arg	Phe	Ile	Gly	Pro	His	Arg	Ile	Val	Ala	Thr	Asn		
	130					135					140						
Asn	Lys	Gly	Lys	Glu	Lys	Ile	Tyr	Ser	Ala	Glu	Arg	Phe	Leu	Ile	Ala		
145				150					155						160		
Thr	Gly	Glu	Arg	Pro	Arg	Tyr	Leu	Gly	Ile	Pro	Gly	Asp	Lys	Glu	Tyr		
			165					170						175			
Cys	Ile	Ser	Ser	Asp	Asp	Leu	Phe	Ser	Leu	Pro	Tyr	Cys	Pro	Gly	Lys		
			180					185					190				
Thr	Leu	Val	Val	Gly	Ala	Ser	Tyr	Val	Ala	Leu	Glu	Cys	Ala	Gly	Phe		
	195						200					205					
Leu	Ala	Gly	Ile	Gly	Leu	Asp	Val	Thr	Val	Met	Val	Arg	Ser	Ile	Leu		
	210					215					220						
Leu	Arg	Gly	Phe	Asp	Gln	Asp	Met	Ala	Asn	Lys	Ile	Gly	Glu	His	Met		
225				230						235					240		
Glu	Glu	His	Gly	Ile	Lys	Phe	Ile	Arg	Gln	Phe	Val	Pro	Thr	Lys	Ile		
			245						250					255			
Glu	Gln	Ile	Glu	Ala	Gly	Thr	Pro	Gly	Arg	Leu	Arg	Val	Thr	Ala	Gln		
			260					265						270			

Ser Thr Asn Ser Glu Glu Thr Ile Glu Gly Glu Phe Asn Thr Val Leu
 275 280 285
 Leu Ala Val Gly Arg Asp Ser Cys Thr Arg Thr Ile Gly Leu Glu Thr
 290 295 300
 Val Gly Val Lys Ile Asn
 305 310

<210> 302
 <211> 613
 <212> PRT
 <213> Mus musculus

<400> 302
 Met Ser Ser Pro Pro Gly Arg Arg Ala Arg Leu Ala Ser Pro Gly Thr
 1 5 10 15
 Ser Arg Pro Ser Ser Glu Ala Arg Glu Glu Leu Arg Arg Arg Leu Arg
 20 25 30
 Asp Leu Ile Glu Gly Asn Arg Val Met Ile Phe Ser Lys Ser Tyr Cys
 35 40 45
 Pro His Ser Thr Arg Val Lys Glu Leu Phe Ser Ser Leu Gly Val Val
 50 55 60
 Tyr Asn Ile Leu Glu Leu Asp Gln Val Asp Asp Gly Ala Ser Val Gln
 65 70 75 80
 Glu Val Leu Thr Glu Ile Ser Asn Gln Lys Thr Val Pro Asn Ile Phe
 85 90 95
 Val Asn Lys Val His Val Gly Gly Cys Asp Arg Thr Phe Gln Ala His
 100 105 110
 Gln Asn Gly Leu Leu Gln Lys Leu Leu Gln Asp Asp Ser Ala His Asp
 115 120 125
 Tyr Asp Leu Ile Ile Ile Gly Gly Gly Ser Gly Gly Leu Ser Cys Ala
 130 135 140
 Lys Glu Ala Ala Asn Leu Gly Lys Lys Val Met Val Leu Asp Phe Val
 145 150 155 160
 Val Pro Ser Pro Gln Gly Thr Thr Trp Gly Leu Gly Gly Thr Cys Val
 165 170 175
 Asn Val Gly Cys Ile Pro Lys Lys Leu Met His Gln Ala Ala Leu Leu
 180 185 190
 Gly His Ala Leu Gln Asp Ala Lys Lys Tyr Gly Trp Glu Tyr Asn Gln
 195 200 205
 Gln Val Lys His Asn Trp Glu Ala Met Thr Glu Ala Ile Gln Ser His
 210 215 220
 Ile Gly Ser Leu Asn Trp Gly Tyr Arg Val Thr Leu Arg Glu Lys Gly
 225 230 235 240
 Val Thr Tyr Val Asn Ser Phe Gly Glu Phe Val Asp Leu His Lys Ile
 245 250 255
 Lys Ala Thr Asn Lys Lys Gly Gln Glu Thr Phe Tyr Thr Ala Ser Lys
 260 265 270
 Phe Val Ile Ala Thr Gly Glu Arg Pro Arg Tyr Leu Gly Ile Gln Gly
 275 280 285
 Asp Lys Glu Tyr Cys Ile Thr Ser Asp Asp Leu Phe Ser Leu Pro Tyr
 290 295 300
 Cys Pro Gly Cys Thr Leu Val Val Gly Ala Ser Tyr Val Gly Leu Glu
 305 310 315 320
 Cys Ala Gly Phe Leu Ala Gly Leu Gly Leu Asp Val Thr Val Met Val
 325 330 335
 Arg Ser Val Leu Leu Arg Gly Phe Asp Gln Glu Met Ala Glu Lys Val
 340 345 350
 Gly Ser Tyr Leu Glu Gln Gln Gly Val Lys Phe Gln Arg Lys Phe Thr
 355 360 365
 Pro Ile Leu Val Gln Gln Leu Glu Lys Gly Leu Pro Gly Lys Leu Lys
 370 375 380
 Val Val Ala Lys Ser Thr Glu Gly Pro Glu Thr Val Glu Gly Ile Tyr
 385 390 395 400
 Asn Thr Val Leu Leu Ala Ile Gly Arg Asp Ser Cys Thr Arg Lys Ile
 405 410 415
 Gly Leu Glu Lys Ile Gly Val Lys Ile Asn Glu Lys Asn Gly Lys Ile

```

      420      425      430
Pro Val Asn Asp Val Glu Gln Thr Asn Val Pro His Val Tyr Ala Ile
      435      440      445
Gly Asp Ile Leu Asp Gly Lys Pro Glu Leu Thr Pro Val Ala Ile Gln
      450      455      460
Ala Gly Lys Leu Leu Ala Arg Arg Leu Phe Gly Val Ser Leu Glu Lys
465      470      475      480
Cys Asp Tyr Ile Asn Ile Pro Thr Thr Val Phe Thr Pro Leu Glu Tyr
      485      490      495
Gly Cys Cys Gly Leu Ser Glu Glu Lys Ala Ile Glu Met Tyr Lys Lys
      500      505      510
Glu Asn Leu Glu Val Tyr His Thr Leu Phe Trp Pro Leu Glu Trp Thr
      515      520      525
Val Ala Gly Arg Asp Asn Asn Thr Cys Tyr Ala Lys Ile Ile Cys Asn
      530      535      540
Lys Phe Asp Asn Glu Arg Val Val Gly Phe His Leu Leu Gly Pro Asn
545      550      555      560
Ala Gly Glu Ile Thr Gln Gly Phe Ala Ala Met Lys Cys Gly Leu
      565      570      575
Thr Lys Gln Leu Leu Asp Asp Thr Ile Gly Ile His Pro Thr Cys Gly
      580      585      590
Glu Val Phe Thr Thr Leu Glu Ile Thr Lys Ser Ser Gly Leu Asp Ile
      595      600      605
Thr Gln Lys Gly Cys
      610

```

<210> 303
 <211> 524
 <212> PRT
 <213> Mus musculus

<220>
 <221> VARIANT
 <222> 523
 <223> Xaa = Any Amino Acid

```

<400> 303
Met Val Ala Ala Met Val Ala Ala Leu Arg Gly Pro Ser Arg Arg Phe
  1      5      10      15
Arg Pro Arg Thr Arg Ala Leu Thr Arg Gly Thr Arg Gly Ala Ala Ser
      20      25      30
Ala Ala Gly Gln Gln Ser Phe Asp Leu Leu Val Ile Gly Gly Gly
      35      40      45
Ser Gly Gly Leu Ala Cys Ala Lys Glu Ala Ala Gln Leu Gly Lys Lys
      50      55      60
Val Ala Val Ala Asp Tyr Val Glu Pro Ser Pro Arg Gly Thr Lys Trp
65      70      75      80
Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys Ile Pro Lys Lys Leu
      85      90      95
Met His Gln Ala Ala Leu Leu Gly Gly Met Ile Arg Asp Ala His His
      100      105      110
Tyr Gly Trp Glu Val Ala Gln Pro Val Gln His Asn Trp Lys Thr Met
      115      120      125
Ala Glu Ala Val Gln Asn His Val Lys Ser Leu Asn Trp Gly His Arg
      130      135      140
Val Gln Leu Gln Asp Arg Lys Val Lys Tyr Phe Asn Ile Lys Ala Ser
145      150      155      160
Phe Val Asp Glu His Thr Val Arg Gly Val Asp Lys Gly Gly Lys Ala
      165      170      175
Thr Leu Leu Ser Ala Glu His Ile Val Ile Ala Thr Gly Gly Arg Pro
      180      185      190
Arg Tyr Pro Thr Gln Val Lys Gly Ala Leu Glu Tyr Gly Ile Thr Ser
      195      200      205
Asp Asp Ile Phe Trp Leu Lys Glu Ser Pro Gly Lys Thr Leu Val Val
      210      215      220
Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe Leu Thr Gly Ile

```

<400> 304															
Met	Ala	Ala	Met	Val	Ala	Gly	Arg	Met	Trp	Ala	Ala	Leu	Arg	Gly	Pro
1				5					10					15	
Ser	Arg	Arg	Phe	Arg	Pro	Arg	Thr	Arg	Ala	Leu	Thr	Arg	Gly	Thr	Arg
			20					25					30		
Gly	Ala	Ala	Ser	Ala	Ala	Gly	Gly	Gln	Gln	Ser	Phe	Asp	Leu	Leu	Val
		35					40					45			
Ile	Gly	Gly	Gly	Ser	Gly	Gly	Leu	Ala	Cys	Ala	Lys	Glu	Ala	Ala	Gln
	50					55					60				
Leu	Gly	Lys	Lys	Val	Ala	Val	Ala	Asp	Tyr	Val	Glu	Pro	Ser	Pro	Arg
65				70					75					80	
Gly	Thr	Lys	Trp	Gly	Leu	Gly	Gly	Thr	Cys	Val	Asn	Val	Gly	Cys	Ile
				85				90						95	
Pro	Lys	Lys	Leu	Met	His	Gln	Ala	Ala	Leu	Leu	Gly	Gly	Met	Ile	Arg
			100					105					110		
Asp	Ala	His	His	Tyr	Gly	Trp	Glu	Val	Ala	Gln	Pro	Val	Gln	His	Asn
		115					120					125			
Trp	Lys	Thr	Met	Ala	Glu	Ala	Val	Gln	Asn	His	Val	Lys	Ser	Leu	Asn

```
<210> 305
<211> 520
<212> PRT
<213> Mus musculus
```

-193-

```

Thr Cys Val Asn Val Gly Cys Ile Pro Lys Lys Leu Met His Gln Ala
      85      90      95
Ala Leu Leu Gly Gly Met Ile Arg Asp Ala His His Tyr Gly Trp Glu
      100      105      110
Val Ala Gln Pro Val Gln His Asn Trp Lys Thr Met Ala Glu Ala Val
      115      120      125
Gln Asn His Val Lys Ser Leu Asn Trp Gly His Arg Val Gln Leu Gln
      130      135      140
Asp Arg Lys Val Lys Tyr Phe Asn Ile Lys Ala Ser Phe Val Asp Glu
      145      150      155      160
His Thr Val Arg Gly Val Asp Lys Gly Gly Lys Ala Thr Leu Leu Ser
      165      170      175
Ala Glu His Ile Val Ile Ala Thr Gly Gly Arg Pro Arg Tyr Pro Thr
      180      185      190
Gln Val Lys Gly Ala Leu Glu Tyr Gly Ile Thr Ser Asp Asp Ile Phe
      195      200      205
Trp Leu Lys Glu Ser Pro Gly Lys Thr Leu Val Val Gly Ala Ser Tyr
      210      215      220
Val Ala Leu Glu Cys Ala Gly Phe Leu Thr Gly Ile Gly Leu Asp Thr
      225      230      235      240
Thr Val Met Met Arg Ser Ile Pro Leu Arg Gly Phe Asp Gln Gln Met
      245      250      255
Ser Ser Leu Val Thr Glu His Met Glu Ser His Gly Thr Gln Phe Leu
      260      265      270
Lys Gly Cys Val Pro Ser His Ile Lys Lys Leu Pro Thr Asn Gln Leu
      275      280      285
Gln Val Thr Trp Glu Asp His Ala Ser Gly Lys Glu Asp Thr Gly Thr
      290      295      300
Phe Asp Thr Val Leu Trp Ala Ile Gly Arg Val Pro Glu Thr Arg Thr
      305      310      315      320
Leu Asn Leu Glu Lys Ala Gly Ile Ser Thr Asn Pro Lys Asn Gln Lys
      325      330      335
Ile Ile Val Asp Ala Gln Glu Ala Thr Ser Val Pro His Ile Tyr Ala
      340      345      350
Ile Gly Asp Val Ala Glu Gly Arg Pro Glu Leu Thr Pro Thr Ala Ile
      355      360      365
Lys Ala Gly Lys Leu Leu Ala Gln Arg Leu Phe Gly Lys Ser Ser Thr
      370      375      380
Leu Met Asp Tyr Ser Asn Val Pro Thr Thr Val Phe Thr Pro Leu Glu
      385      390      395      400
Tyr Gly Cys Val Gly Leu Ser Glu Glu Glu Ala Val Ala Leu His Gly
      405      410      415
Gln Glu His Val Glu Val Tyr His Ala Tyr Tyr Lys Pro Leu Glu Phe
      420      425      430
Thr Val Ala Asp Arg Asp Ala Ser Gln Cys Tyr Ile Lys Met Val Cys
      435      440      445
Met Arg Glu Pro Pro Gln Leu Val Leu Gly Leu His Phe Leu Gly Pro
      450      455      460
Asn Ala Gly Glu Val Thr Gln Gly Phe Ala Leu Gly Ile Lys Cys Gly
      465      470      475      480
Ala Ser Tyr Ala Gln Val Met Gln Thr Val Gly Ile His Pro Thr Cys
      485      490      495
Ser Glu Glu Val Val Lys Leu His Ile Ser Lys Arg Ser Gly Leu Glu
      500      505      510
Pro Thr Val Thr Gly Cys Cys Gly
      515      520

```

<210> 306

<211> 499

<212> PRT

<213> Mus musculus

<400> 306

```

Met Asn Gly Ser Lys Asp Pro Pro Gly Ser Tyr Asp Phe Asp Leu Ile
1      5      10      15
Ile Ile Gly Gly Gly Ser Gly Gly Leu Ala Ala Ala Lys Glu Ala Ala

```

			20					25					30				
Lys	Phe	Asp	Lys	Lys	Val	Leu	Val	Leu	Asp	Phe	Val	Thr	Pro	Thr	Pro		
		35					40					45					
Leu	Gly	Thr	Arg	Trp	Gly	Leu	Gly	Gly	Thr	Cys	Val	Asn	Val	Gly	Cys		
	50					55					60						
Ile	Pro	Lys	Lys	Leu	Met	His	Gln	Ala	Ala	Leu	Gly	Gln	Ala	Leu			
65					70					75				80			
Lys	Asp	Ser	Arg	Asn	Tyr	Gly	Trp	Lys	Val	Glu	Asp	Thr	Val	Lys	His		
				85					90					95			
Asp	Trp	Glu	Lys	Met	Thr	Glu	Ser	Val	Gln	Ser	His	Ile	Gly	Ser	Leu		
			100					105					110				
Asn	Trp	Gly	Tyr	Arg	Val	Ala	Leu	Arg	Glu	Lys	Lys	Val	Val	Tyr	Glu		
		115					120					125					
Asn	Ala	Tyr	Gly	Arg	Phe	Ile	Gly	Pro	His	Arg	Ile	Val	Ala	Thr	Asn		
	130					135					140						
Asn	Lys	Gly	Lys	Glu	Lys	Ile	Tyr	Ser	Ala	Glu	Arg	Phe	Leu	Ile	Ala		
145					150					155				160			
Thr	Gly	Glu	Arg	Pro	Arg	Tyr	Leu	Gly	Ile	Pro	Gly	Asp	Lys	Glu	Tyr		
				165				170						175			
Cys	Ile	Ser	Ser	Asp	Asp	Leu	Phe	Ser	Leu	Pro	Tyr	Cys	Pro	Gly	Lys		
			180					185					190				
Thr	Leu	Val	Val	Gly	Ala	Ser	Tyr	Val	Ala	Leu	Glu	Cys	Ala	Gly	Phe		
	195						200					205					
Leu	Ala	Gly	Ile	Gly	Leu	Asp	Val	Thr	Val	Met	Val	Arg	Ser	Ile	Leu		
	210					215					220						
Leu	Arg	Gly	Phe	Asp	Gln	Asp	Met	Ala	Asn	Lys	Ile	Gly	Glu	His	Met		
225					230					235				240			
Glu	Glu	His	Gly	Ile	Lys	Phe	Ile	Arg	Gln	Phe	Val	Pro	Thr	Lys	Ile		
				245					250					255			
Glu	Gln	Ile	Glu	Ala	Gly	Thr	Pro	Gly	Arg	Leu	Arg	Val	Thr	Ala	Gln		
			260					265					270				
Ser	Thr	Asn	Ser	Glu	Glu	Thr	Ile	Glu	Gly	Glu	Phe	Asn	Thr	Val	Leu		
		275					280					285					
Leu	Ala	Val	Gly	Arg	Asp	Ser	Cys	Thr	Arg	Thr	Ile	Gly	Leu	Glu	Thr		
	290					295					300						
Val	Gly	Val	Lys	Ile	Asn	Glu	Lys	Thr	Gly	Lys	Ile	Pro	Val	Thr	Asp		
305					310					315					320		
Glu	Glu	Gln	Thr	Asn	Val	Pro	Tyr	Ile	Tyr	Ala	Ile	Gly	Asp	Ile	Leu		
				325					330					335			
Glu	Gly	Lys	Leu	Glu	Leu	Thr	Pro	Val	Ala	Ile	Gln	Ala	Gly	Arg	Leu		
			340					345					350				
Leu	Ala	Gln	Arg	Leu	Tyr	Gly	Gly	Ser	Asn	Val	Lys	Cys	Asp	Tyr	Asp		
	355						360					365					
Asn	Val	Pro	Thr	Thr	Val	Phe	Thr	Pro	Leu	Glu	Tyr	Gly	Cys	Cys	Gly		
	370					375					380						
Leu	Ser	Glu	Glu	Lys	Ala	Val	Glu	Lys	Phe	Gly	Glu	Glu	Asn	Ile	Glu		
385					390					395				400			
Val	Tyr	His	Ser	Phe	Phe	Trp	Pro	Leu	Glu	Trp	Thr	Val	Pro	Ser	Arg		
				405					410					415			
Asp	Asn	Asn	Lys	Cys	Tyr	Ala	Lys	Ile	Ile	Cys	Asn	Leu	Lys	Asp	Asp		
			420					425					430				
Glu	Arg	Val	Val	Gly	Phe	His	Val	Leu	Gly	Pro	Asn	Ala	Gly	Glu	Val		
		435					440					445					
Thr	Gln	Gly	Phe	Ala	Ala	Ala	Leu	Lys	Cys	Gly	Leu	Thr	Lys	Gln	Gln		
	450					455					460						
Leu	Asp	Ser	Thr	Ile	Gly	Ile	His	Pro	Val	Cys	Ala	Glu	Ile	Phe	Thr		
465					470					475				480			
Thr	Leu	Ser	Val	Thr	Lys	Arg	Ser	Gly	Gly	Asp	Ile	Leu	Gln	Ser	Gly		
				485					490					495			
Cys	Cys	Gly															

<210> 307

<211> 497

<212> PRT

<213> Rattus norvegicus

<220>

<221> VARIANT

<222> 497

<223> Xaa = Any Amino Acid

<400> 307

Met	Asn	Asp	Ser	Lys	Asp	Ala	Pro	Lys	Ser	Tyr	Asp	Phe	Asp	Leu	Ile
1				5					10					15	
Ile	Ile	Gly	Gly	Gly	Ser	Gly	Gly	Leu	Ala	Ala	Ala	Lys	Glu	Ala	Ala
		20						25					30		
Lys	Phe	Asp	Lys	Lys	Val	Met	Val	Leu	Asp	Phe	Val	Thr	Pro	Thr	Pro
		35					40					45			
Leu	Gly	Thr	Asn	Gly	Gly	Leu	Gly	Gly	Thr	Cys	Val	Asn	Val	Gly	Cys
	50					55					60				
Ile	Pro	Lys	Lys	Leu	Met	His	Gln	Ala	Ala	Leu	Leu	Gly	Gln	Ala	Leu
65					70					75					80
Lys	Asp	Ser	Arg	Asn	Tyr	Gly	Trp	Lys	Leu	Glu	Asp	Thr	Val	Lys	His
				85					90					95	
Asp	Trp	Glu	Lys	Met	Thr	Glu	Ser	Val	Gln	Asn	His	Ile	Gly	Ser	Leu
			100					105					110		
Asn	Trp	Gly	Tyr	Arg	Val	Ala	Leu	Arg	Glu	Lys	Lys	Val	Val	Tyr	Glu
		115					120					125			
Asn	Ala	Tyr	Gly	Lys	Phe	Ile	Gly	Pro	His	Lys	Ile	Met	Ala	Thr	Asn
	130					135					140				
Asn	Lys	Gly	Lys	Glu	Lys	Val	Tyr	Ser	Ala	Glu	Arg	Phe	Leu	Ile	Ala
145					150					155					160
Thr	Gly	Glu	Arg	Pro	Arg	Tyr	Leu	Gly	Ile	Pro	Gly	Asp	Lys	Glu	Tyr
				165				170						175	
Cys	Ile	Ser	Ser	Asp	Asp	Leu	Phe	Ser	Leu	Pro	Tyr	Cys	Pro	Gly	Lys
			180					185					190		
Thr	Leu	Val	Val	Gly	Ala	Ser	Tyr	Val	Ala	Leu	Glu	Cys	Ala	Gly	Phe
		195					200					205			
Leu	Ala	Gly	Ile	Gly	Leu	Asp	Val	Thr	Val	Met	Val	Arg	Ser	Ile	Leu
	210					215					220				
Leu	Arg	Gly	Phe	Asp	Gln	Asp	Met	Ala	Asn	Lys	Ile	Gly	Glu	His	Met
225					230					235					240
Glu	Glu	His	Gly	Ile	Lys	Phe	Ile	Arg	Gln	Phe	Val	Pro	Thr	Lys	Ile
				245					250					255	
Glu	Gln	Ile	Glu	Ala	Gly	Thr	Pro	Gly	Arg	Leu	Lys	Val	Thr	Ala	Lys
			260					265					270		
Ser	Thr	Asn	Ser	Glu	Glu	Thr	Ile	Glu	Asp	Glu	Phe	Asn	Thr	Val	Leu
		275					280					285			
Leu	Ala	Val	Gly	Arg	Asp	Ser	Cys	Thr	Arg	Thr	Ile	Gly	Leu	Glu	Thr
	290					295					300				
Val	Gly	Val	Lys	Ile	Asn	Glu	Lys	Thr	Gly	Lys	Ile	Pro	Val	Thr	Asp
305					310					315					320
Glu	Glu	Gln	Thr	Asn	Val	Pro	Tyr	Ile	Tyr	Ala	Ile	Gly	Asp	Ile	Leu
				325					330					335	
Glu	Gly	Lys	Leu	Glu	Leu	Thr	Pro	Val	Ala	Ile	Gln	Ala	Gly	Arg	Leu
			340					345					350		
Leu	Ala	Gln	Arg	Leu	Tyr	Gly	Gly	Ser	Thr	Val	Lys	Cys	Asp	Tyr	Asp
		355					360					365			
Asn	Val	Pro	Thr	Thr	Val	Phe	Thr	Pro	Leu	Glu	Tyr	Gly	Cys	Cys	Gly
	370					375					380				
Leu	Ser	Glu	Glu	Lys	Ala	Val	Glu	Lys	Phe	Gly	Glu	Glu	Asn	Ile	Glu
385					390					395					400
Val	Tyr	His	Ser	Phe	Trp	Pro	Leu	Glu	Trp	Thr	Val	Pro	Ser	Arg	
				405				410					415		
Asp	Asn	Asn	Lys	Cys	Tyr	Ala	Lys	Val	Ile	Cys	Asn	Leu	Lys	Asp	Asn
			420					425					430		
Glu	Arg	Val	Val	Gly	Phe	His	Val	Leu	Gly	Pro	Asn	Ala	Gly	Glu	Val
		435					440					445			
Thr	Gln	Gly	Phe	Ala	Ala	Ala	Leu	Lys	Cys	Gly	Leu	Thr	Lys	Gln	Gln
	450					455					460				
Leu	Asp	Ser	Thr	Ile	Gly	Ile	His	Pro	Val	Cys	Ala	Glu	Ile	Phe	Thr
465					470					475					480
Thr	Leu	Ser	Val	Thr	Lys	Arg	Ser	Gly	Gly	Asp	Ile	Leu	Gln	Ser	Gly

```
<210> 309
<211> 498
<212> PRT
<213> Rattus norvegicus
```

<400>	309															
Met	Asn	Asp	Ser	Lys	Asp	Ala	Pro	Lys	Ser	Tyr	Asp	Phe	Asp	Leu	Ile	
1				5					10					15		
Ile	Ile	Gly	Gly	Gly	Ser	Gly	Gly	Leu	Ala	Ala	Ala	Lys	Glu	Ala	Ala	
			20					25					30			
Lys	Phe	Asp	Lys	Lys	Val	Met	Val	Leu	Asp	Phe	Val	Thr	Pro	Thr	Pro	
		35					40					45				
Leu	Gly	Thr	Asn	Gly	Gly	Leu	Gly	Gly	Thr	Cys	Val	Asn	Val	Gly	Cys	
		50				55					60					
Ile	Pro	Lys	Lys	Leu	Met	His	Gln	Ala	Ala	Leu	Leu	Gly	Gln	Ala	Leu	
65					70					75					80	
Lys	Asp	Ser	Arg	Asn	Tyr	Gly	Trp	Lys	Leu	Glu	Asp	Thr	Val	Lys	His	
				85					90					95		
Asp	Trp	Glu	Lys	Met	Thr	Glu	Ser	Val	Gln	Asn	His	Ile	Gly	Ser	Leu	
			100					105					110			
Asn	Trp	Gly	Tyr	Arg	Val	Ala	Leu	Arg	Glu	Lys	Lys	Val	Val	Tyr	Glu	
		115					120					125				
Asn	Ala	Tyr	Gly	Lys	Phe	Ile	Gly	Pro	His	Lys	Ile	Met	Ala	Thr	Asn	
		130				135					140					
Asn	Lys	Gly	Lys	Glu	Lys	Val	Tyr	Ser	Ala	Glu	Arg	Phe	Leu	Ile	Ala	
145					150					155					160	
Thr	Gly	Glu	Arg	Pro	Arg	Tyr	Leu	Gly	Ile	Pro	Gly	Asp	Lys	Glu	Tyr	

165 170 175
 Cys Ile Ser Ser Asp Asp Leu Phe Ser Leu Pro Tyr Cys Pro Gly Lys
 180 185 190
 Thr Leu Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe
 195 200 205
 Leu Ala Gly Ile Gly Leu Asp Val Thr Val Met Val Arg Ser Ile Leu
 210 215 220
 Leu Arg Gly Phe Asp Gln Asp Met Ala Asn Lys Ile Gly Glu His Met
 225 230 235 240
 Glu Glu His Gly Ile Lys Phe Ile Arg Gln Phe Val Pro Thr Lys Ile
 245 250 255
 Glu Gln Ile Glu Ala Gly Thr Pro Gly Arg Leu Lys Val Thr Ala Lys
 260 265 270
 Ser Thr Asn Ser Glu Glu Thr Ile Glu Asp Glu Phe Asn Thr Val Leu
 275 280 285
 Leu Ala Val Gly Arg Asp Ser Cys Thr Arg Thr Ile Gly Leu Glu Thr
 290 295 300
 Val Gly Val Lys Ile Asn Glu Lys Thr Gly Lys Ile Pro Val Thr Asp
 305 310 315 320
 Glu Glu Gln Thr Asn Val Pro Tyr Ile Tyr Ala Ile Gly Asp Ile Leu
 325 330 335
 Glu Gly Lys Leu Glu Leu Thr Pro Val Ala Ile Gln Ala Gly Arg Leu
 340 345 350
 Leu Ala Gln Arg Leu Tyr Gly Gly Ser Thr Val Lys Cys Asp Tyr Asp
 355 360 365
 Asn Val Pro Thr Thr Val Phe Thr Pro Leu Glu Tyr Gly Cys Cys Gly
 370 375 380
 Leu Ser Glu Glu Lys Ala Val Glu Lys Phe Gly Glu Glu Asn Ile Glu
 385 390 395 400
 Val Tyr His Ser Phe Phe Trp Pro Leu Glu Trp Thr Val Pro Ser Arg
 405 410 415
 Asp Asn Asn Lys Cys Tyr Ala Lys Val Ile Cys Asn Leu Lys Asp Asn
 420 425 430
 Glu Arg Val Val Gly Phe His Val Leu Gly Pro Asn Ala Gly Glu Val
 435 440 445
 Thr Gln Ala Leu Gln Pro Leu Lys Cys Gly Leu Thr Lys Gln Gln Leu
 450 455 460
 Asp Ser Thr Ile Gly Ile His Pro Val Cys Ala Glu Ile Phe Thr Thr
 465 470 475 480
 Leu Ser Val Thr Lys Arg Ser Gly Gly Asp Ile Leu Gln Ser Gly Cys
 485 490 495
 Xaa Gly

<210> 310
 <211> 11
 <212> PRT
 <213> Rattus norvegicus

<400> 310
 Met Asn Asp Ser Lys Asp Ala Pro Lys Ser Tyr
 1 5 10

<210> 311
 <211> 496
 <212> PRT
 <213> Rattus norvegicus

<400> 311
 Met Asn Asp Ser Lys Asp Ala Pro Lys Ser Tyr Asp Phe Asp Leu Ile
 1 5 10 15
 Ile Ile Gly Gly Gly Ser Gly Gly Leu Ala Ala Ala Lys Glu Ala Ala
 20 25 30
 Lys Phe Asp Lys Lys Val Met Val Leu Asp Phe Val Thr Pro Thr Pro
 35 40 45

Leu Gly Thr Asn Gly Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys
 50 55 60
 Ile Pro Lys Lys Leu Met His Gln Ala Ala Leu Leu Gly Gln Ala Leu
 65 70 75 80
 Lys Asp Ser Arg Asn Tyr Gly Trp Lys Leu Glu Asp Thr Val Lys His
 85 90 95
 Asp Trp Glu Lys Met Thr Glu Ser Val Gln Asn His Ile Gly Ser Leu
 100 105 110
 Asn Trp Gly Tyr Arg Val Ala Leu Arg Glu Lys Lys Val Val Tyr Glu
 115 120 125
 Asn Ala Tyr Gly Lys Phe Ile Gly Pro His Lys Ile Met Ala Thr Asn
 130 135 140
 Asn Lys Gly Lys Glu Lys Val Tyr Ser Ala Glu Arg Phe Leu Ile Ala
 145 150 155 160
 Thr Gly Glu Arg Pro Arg Tyr Leu Gly Ile Pro Gly Asp Lys Glu Tyr
 165 170 175
 Cys Ile Ser Ser Asp Asp Leu Phe Ser Leu Pro Tyr Cys Pro Gly Lys
 180 185 190
 Thr Leu Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe
 195 200 205
 Leu Ala Gly Ile Gly Leu Asp Val Thr Val Met Val Arg Ser Ile Leu
 210 215 220
 Leu Arg Gly Phe Asp Gln Asp Met Ala Asn Lys Ile Gly Glu His Met
 225 230 235 240
 Glu Glu His Gly Ile Lys Phe Ile Arg Gln Phe Val Pro Thr Lys Ile
 245 250 255
 Glu Gln Ile Glu Ala Gly Thr Pro Gly Arg Leu Lys Val Thr Ala Lys
 260 265 270
 Ser Thr Asn Ser Glu Glu Thr Ile Glu Asp Glu Phe Asn Thr Val Leu
 275 280 285
 Leu Ala Val Gly Arg Asp Ser Cys Thr Arg Thr Ile Gly Leu Glu Thr
 290 295 300
 Val Gly Val Lys Ile Asn Glu Lys Thr Gly Lys Ile Pro Val Thr Asp
 305 310 315 320
 Glu Glu Gln Thr Asn Val Pro Tyr Ile Tyr Ala Ile Gly Asp Ile Leu
 325 330 335
 Glu Gly Lys Leu Glu Leu Thr Pro Val Ala Ile Gln Ala Gly Arg Leu
 340 345 350
 Leu Ala Gln Arg Leu Tyr Gly Gly Ser Thr Val Lys Cys Asp Tyr Asp
 355 360 365
 Asn Val Pro Thr Thr Val Phe Thr Pro Leu Glu Tyr Gly Cys Cys Gly
 370 375 380
 Leu Ser Glu Glu Lys Ala Val Glu Lys Phe Gly Glu Glu Asn Ile Glu
 385 390 395 400
 Val Tyr His Ser Phe Phe Trp Pro Leu Glu Trp Thr Val Pro Ser Arg
 405 410 415
 Asp Asn Asn Lys Cys Tyr Ala Lys Val Ile Cys Asn Leu Lys Asp Asn
 420 425 430
 Glu Arg Val Val Gly Phe His Val Leu Gly Pro Asn Ala Gly Glu Val
 435 440 445
 Thr Gln Ala Leu Gln Pro Leu Lys Cys Gly Leu Thr Lys Gln Gln Leu
 450 455 460
 Asp Ser Thr Ile Gly Ile His Pro Val Cys Ala Glu Ile Phe Thr Thr
 465 470 475 480
 Leu Ser Val Thr Lys Arg Ser Gly Gly Asp Ile Leu Gln Ser Gly Cys
 485 490 495

<210> 312

<211> 526

<212> PRT

<213> Rattus norvegicus

<220>

<221> VARIANT

<222> 525

<223> Xaa = Any Amino Acid

<400> 312
 Met Ala Ala Ile Val Ala Ala Leu Arg Gly Ser Ser Gly Arg Phe Arg
 1 5 10 15
 Pro Gln Thr Arg Val Leu Thr Arg Gly Thr Arg Gly Ala Ala Gly Ala
 20 25 30
 Ala Ser Ala Ala Gly Gly Gln Gln Asn Phe Asp Leu Leu Val Ile Gly
 35 40 45
 Gly Gly Ser Gly Gly Leu Ala Cys Ala Lys Glu Ala Ala Gln Leu Gly
 50 55 60
 Arg Lys Val Ala Val Ala Asp Tyr Val Glu Pro Ser Pro Arg Gly Thr
 65 70 75 80
 Lys Trp Gly Leu Gly Gly Thr Cys Val Asn Val Gly Cys Ile Pro Lys
 85 90 95
 Lys Leu Met His Gln Ala Ala Leu Leu Gly Gly Met Ile Arg Asp Ala
 100 105 110
 Gln His Tyr Gly Trp Glu Val Ala Gln Pro Val Gln His Asn Trp Lys
 115 120 125
 Ala Met Ala Glu Ala Val Gln Asn His Val Lys Ser Leu Asn Trp Gly
 130 135 140
 His Arg Val Gln Leu Gln Asp Arg Lys Val Lys Tyr Phe Asn Ile Lys
 145 150 155 160
 Ala Ser Phe Val Asn Glu His Thr Val His Gly Val Asp Lys Ala Gly
 165 170 175
 Lys Val Thr Gln Leu Ser Ala Lys His Ile Val Ile Ala Thr Gly Gly
 180 185 190
 Arg Pro Lys Tyr Pro Thr Gln Val Lys Gly Ala Leu Glu His Gly Ile
 195 200 205
 Thr Ser Asp Asp Ile Phe Trp Leu Lys Glu Ser Pro Gly Lys Thr Leu
 210 215 220
 Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe Leu Thr
 225 230 235 240
 Gly Ile Gly Leu Asp Thr Thr Val Met Met Arg Ser Val Pro Leu Arg
 245 250 255
 Gly Phe Asp Gln Gln Met Ala Ser Leu Val Thr Glu His Met Glu Ser
 260 265 270
 His Gly Thr Arg Phe Leu Lys Gly Cys Val Pro Ser Leu Ile Arg Lys
 275 280 285
 Leu Pro Thr Asn Gln Leu Gln Val Thr Trp Glu Asp Leu Ala Ser Gly
 290 295 300
 Lys Glu Asp Val Gly Thr Phe Asp Thr Val Leu Trp Ala Ile Gly Arg
 305 310 315 320
 Val Pro Glu Thr Arg Asn Leu Asn Leu Glu Lys Ala Gly Val Asn Thr
 325 330 335
 Asn Pro Lys Asn Gln Lys Ile Ile Val Asp Ala Gln Glu Ala Thr Ser
 340 345 350
 Val Pro His Ile Tyr Ala Ile Gly Asp Val Ala Glu Gly Arg Pro Glu
 355 360 365
 Leu Thr Pro Thr Ala Ile Lys Ala Gly Lys Leu Leu Ala Gln Arg Leu
 370 375 380
 Phe Gly Lys Ser Ser Thr Leu Met Asn Tyr Ser Asn Val Pro Thr Thr
 385 390 395 400
 Val Phe Thr Pro Leu Glu Tyr Gly Cys Val Gly Leu Ser Glu Glu Glu
 405 410 415
 Ala Val Ala Leu His Gly Gln Glu His Ile Glu Val Tyr His Ala Tyr
 420 425 430
 Tyr Lys Pro Leu Glu Phe Thr Val Ala Asp Arg Asp Ala Ser Gln Cys
 435 440 445
 Tyr Ile Lys Met Val Cys Met Arg Glu Pro Pro Gln Leu Val Leu Gly
 450 455 460
 Leu His Phe Leu Gly Pro Asn Ala Gly Glu Val Thr Gln Gly Phe Ala
 465 470 475 480
 Leu Gly Ile Gln Cys Gly Ala Ser Tyr Ala Gln Val Met Gln Thr Val
 485 490 495
 Gly Ile His Pro Thr Cys Ser Glu Glu Val Val Lys Leu His Ile Ser
 500 505 510
 Lys Arg Ser Gly Leu Asp Pro Thr Val Thr Gly Cys Xaa Gly
 515 520 525

<210> 313
 <211> 499
 <212> PRT
 <213> Sus Scrofa

<220>
 <221> VARIANT
 <222> 498
 <223> Xaa = Any Amino Acid

<400> 313

```

Met Asn Gly Ala Glu Glu Leu Pro Glu Met Tyr Asp Tyr Asp Leu Ile
 1          5          10          15
Ile Ile Gly Gly Ser Gly Gly Leu Ala Ala Ala Lys Glu Ala Ala
 20          25          30
Arg Phe Asn Lys Arg Val Met Val Leu Asp Phe Val Thr Pro Thr Pro
 35          40          45
Leu Gly Thr Arg Trp Gly Leu Gly Gly Thr Cys Val Asn Val Ser Cys
 50          55          60
Ile Pro Lys Lys Leu Met His Gln Ala Ala Leu Leu Gly Gln Ala Leu
 65          70          75          80
Arg Asp Ser Arg Asn Tyr Gly Trp Asn Val Glu Glu Thr Ile Lys His
 85          90          95
Asp Trp Glu Arg Met Thr Glu Ala Val Gln Asn His Ile Gly Ser Leu
 100         105         110
Asn Trp Gly Tyr Arg Val Ala Leu Arg Glu Lys Lys Val Thr Tyr Glu
 115         120         125
Asn Ala Tyr Gly Gln Phe Val Gly Pro His Arg Ile Lys Ala Thr Asn
 130         135         140
Asn Lys Gly Lys Glu Lys Ile Tyr Ser Ala Glu Lys Phe Leu Ile Ala
 145         150         155         160
Thr Gly Glu Arg Pro Arg Tyr Leu Gly Ile Pro Gly Asp Lys Glu Tyr
 165         170         175
Cys Ile Ser Ser Asp Asp Leu Phe Ser Leu Pro Tyr Cys Pro Gly Lys
 180         185         190
Thr Leu Val Val Gly Ala Ser Tyr Val Ala Leu Glu Cys Ala Gly Phe
 195         200         205
Leu Ala Gly Ile Gly Leu Asp Val Thr Val Met Val Arg Ser Ile Leu
 210         215         220
Leu Arg Gly Phe Asp Gln Asp Met Ala Asn Lys Ile Gly Glu His Met
 225         230         235         240
Glu Glu His Gly Ile Lys Phe Ile Arg Gln Phe Val Pro Ile Lys Val
 245         250         255
Glu Gln Ile Glu Ala Gly Thr Pro Gly Arg Leu Arg Val Val Ala Gln
 260         265         270
Ser Thr Asn Ser Glu Glu Ile Ile Glu Gly Glu Tyr Asn Thr Val Met
 275         280         285
Leu Ala Ile Gly Arg Asp Ala Cys Thr Arg Lys Ile Gly Leu Glu Thr
 290         295         300
Val Gly Val Lys Ile Asn Glu Lys Thr Gly Lys Ile Pro Val Thr Asp
 305         310         315         320
Glu Glu Gln Thr Asn Val Pro Tyr Ile Tyr Ala Ile Gly Asp Ile Leu
 325         330         335
Glu Asp Lys Val Glu Leu Thr Pro Val Ala Ile Gln Ala Gly Arg Leu
 340         345         350
Leu Ala Gln Arg Leu Tyr Ala Gly Ser Thr Val Lys Cys Asp Tyr Glu
 355         360         365
Asn Val Pro Thr Thr Val Phe Thr Pro Leu Glu Tyr Gly Ala Cys Gly
 370         375         380
Leu Ser Glu Glu Lys Ala Val Glu Lys Phe Gly Glu Glu Asn Ile Glu
 385         390         395         400
Val Tyr His Ser Tyr Phe Trp Pro Leu Glu Trp Thr Ile Pro Ser Arg
 405         410         415
Asp Asn Asn Lys Cys Tyr Ala Lys Ile Ile Cys Asn Thr Lys Asp Asn
 420         425         430
Glu Arg Val Val Gly Phe His Val Leu Gly Pro Asn Ala Gly Glu Val

```

		435					440					445							
Thr	Gln	Gly	Phe	Ala	Ala	Ala	Leu	Lys	Cys	Gly	Leu	Thr	Lys	Lys	Gln				
	450						455					460							
Leu	Asp	Ser	Thr	Ile	Gly	Ile	His	Pro	Val	Cys	Ala	Glu	Val	Phe	Thr				
465					470					475					480				
Thr	Leu	Ser	Val	Thr	Lys	Arg	Ser	Gly	Ala	Ser	Ile	Leu	Gln	Ala	Gly				
				485					490					495					
Cys	Xaa	Gly																	

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/50240

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/79

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 53698 A (BOOTHE JOSEPH ;DECKERS HARM M (CA); GOLL JANIS (CA); MOLONEY MAURI) 3 December 1998 (1998-12-03) page 13, line 27 -page 14, line 4; claims 1-42; example 11	1-266
X	US 5 948 682 A (MOLONEY, MAURICE M.) 7 September 1999 (1999-09-07) column 18, line 46 - line 55	1-266

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

Date of the actual completion of the international search

8 May 2002

Date of mailing of the international search report

21/05/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Sprinks, M

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/50240

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9853698	A	03-12-1998	AU 737896 B2	06-09-2001
			AU 7517898 A	30-12-1998
			BR 9809691 A	03-10-2000
			WO 9853698 A1	03-12-1998
			CN 1258198 T	28-06-2000
			EP 0986309 A1	22-03-2000
			JP 2002503268 T	29-01-2002
			NO 995802 A	26-01-2000
			US 6146645 A	14-11-2000
			US 6183762 B1	06-02-2001
			US 6210742 B1	03-04-2001
			US 2002037303 A1	28-03-2002
			ZA 9804459 A	13-04-1999
US 5948682	A	07-09-1999	US 5650554 A	22-07-1997
			US 6288304 B1	11-09-2001
			AU 709141 B2	19-08-1999
			AU 4295096 A	24-07-1996
			BR 9600006 A	21-01-1998
			CA 2208751 A1	11-07-1996
			WO 9621029 A1	11-07-1996
			EP 0871749 A1	21-10-1998
			ZA 9510999 A	13-07-1996
			WO 9321320 A1	28-10-1993
			WO 9320216 A1	14-10-1993